CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Applications of metabolomics to study the pathophysiology of adverse pregnancy outcomes
Authors	Morillon, Aude-Claire
Publication date	2020
Original Citation	Morillon, A-C. 2020. Applications of metabolomics to study the pathophysiology of adverse pregnancy outcomes. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2020, Aude-Claire Morillon.
Download date	2026-11-17 11:51:39
Item downloaded from	https://hdl.handle.net/10468/11933



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hÉireann, Corcaigh

National University of Ireland, Cork



Applications of Metabolomics to Study the Pathophysiology of Adverse Pregnancy Outcomes

Thesis presented by

Aude-Claire Morillon, MSc

ORCID 0000-0002-9692-9396

for the degree of

Doctor of Philosophy

University College Cork

Department of Obstetrics and Gynaecology

Head of department: Prof John Higgins

Supervisors: Dr Fergus McCarthy, Dr Jane English,

Dr Shirish Yakkundi, Prof. Louise Kenny, Prof. Philip Baker

2020

Table of Contents

Table of Contents ..... **i**

Declaration ..... **viii**

Acknowledgments ..... **ix**

Thesis Abstract ..... **x**

Contribution and List of Publications ..... **xv**

Abbreviations list ..... **xvii**

Chapter 1: **1**

Introduction **1**

1.1. Small for gestational age 2

1.1.1. Definition and incidence 2

1.1.2. Impact on health 3

1.1.3. Aetiology of Small for Gestational Age 4

1.1.3.1. Genetic factors 4

1.1.3.2. Maternal factors 7

1.1.3.3. Fetal factors 8

1.1.3.4. Placental factors 9

1.1.4. Predictors of Small for Gestational Age 11

1.2. Preterm Birth 13

1.2.1. Definition and incidence 13

1.2.2. Impact on health 14

1.2.3. Aetiology of Preterm Birth 14

1.2.3.1. Infection/Inflammation 15

1.2.3.2. Uteroplacental ischemia 16

1.2.3.3. Uterine overdistension 17

1.2.3.4. Cervical disorders 18
1.2.3.5. Maternal stress 19
1.2.3.6. Endocrine disorders 19
1.2.3.7. Risk factors associated with Preterm Birth 20
1.2.4. Predictors of Preterm Birth 22
1.3. Pre-eclampsia 24
1.3.1. Definition and incidence 24
1.3.2. Impact on health 24
1.3.3. Aetiology of Pre-eclampsia 25
1.3.4. Risk factors associated with Pre-eclampsia 29
1.3.5. Predictors of Pre-eclampsia 31
1.4. Metabolomics 34
1.4.1. Metabolites, metabolome and metabolomics 34
1.4.2. Metabolomics analytical platforms 36
1.4.2.1. Mass spectrometry 36
1.4.2.2. Liquid chromatography mass spectrometry (LC-MS) 38
1.4.2.3. Other analytical platforms frequently used in metabolomics 39
1.4.3. Metabolomics workflow 41
1.4.3.1. Sample collection, sample preparation and data acquisition 42
1.4.3.2. Bioinformatics for metabolomics 43
1.4.4. Workflow for clinical metabolomics 46
1.4.4.1. Biomarker discovery or untargeted metabolic profiling 47
1.4.4.2. Targeted analysis 49
1.4.5. Challenges in clinical metabolomics 52
1.4.5.1. Challenges with sample collection and storage 52
1.4.5.2. Challenges with sample size and power 53

1.4.5.3. Challenges with biomarker validation 54

1.5. Hypotheses and aims 56

Chapter 2: 58

Examining the predictive accuracy of metabolomics for small for gestational age babies: a systematic review. 58

2.1. Abstract 59

2.2. Introduction 60

2.3. Methods 61

2.3.1. Literature search strategy 61

2.3.2. Outcomes and subgroup analysis 64

2.3.3. Selection criteria of studies, data collection and analysis 64

2.3.4. Risk of bias and applicability concerns assessment concerns regarding applicability 65

2.3.5. Data synthesis 65

2.3.6. Patient and public involvement 66

2.4. Results 66

2.4.1. Literature search characteristics 66

2.4.2. Characteristics of the included studies 67

2.4.3. Subgroup analysis 73

2.4.4. Analysis of identified metabolites 79

2.4.5. Risk of bias and Applicability Concerns 82

2.4.6. Meta-analysis 84

2.5. Discussion 85

2.5.1. Main findings 85

2.5.2. Sub group findings 87

2.5.3. Methodological quality 89

2.5.4. Strengths and limitations 89
2.6. Conclusions and implications for practice 90
Chapter 3: 91
Glycerophospholipid and Detoxification Pathways Associated with Small for Gestation Age Pathophysiology: Discovery Metabolomics Analysis in the SCOPE Cohort. 91
3.1. Abstract 92
3.2. Introduction 93
3.3. Materials and Methods 94
3.3.1. Participants 94
3.3.2. Reagents and materials 95
3.3.3. Sample preparations 95
3.3.3.1. Plasma samples 95
3.3.3.2. Urine samples 96
3.3.4. LC analysis 96
3.3.4.1. Plasma samples 96
3.3.4.2. Urine samples 97
3.3.5. MS configuration 98
3.3.6. Data processing 98
3.3.7. statistical analysis 98
3.3.8. Metabolites putative annotation 101
3.4. Results 101
3.4.1. Global lipidomic analysis of plasma samples 101
3.4.2. Global metabolomics analysis of urine samples 108
3.5. Discussion and conclusions 111
Chapter 4: 115

Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomics analysis in the Cork pregnancy cohort. ... 115

4.1. Abstract 116

4.2. Introduction....117

4.3. Materials and Methods 119

4.3.1. Study participants 119

4.3.2. Reagents and materials....123

4.3.3. Sample preparation....123

4.3.4. Global Lipidomics profiling analysis 123

4.3.5. Data pre-processing 124

4.3.6. Statistical analysis 125

4.3.7. Putative annotation of lipids....126

4.4. Results 126

4.5. Discussion 130

Chapter 5:..... 133

Altered Plasma Levels of Phosphatidylcholine and Triacylglycerol in Preterm Birth; A Biomarker Discovery and Validation study in the SCOPE Birth Cohort..... 133

5.1. Abstract 134

5.2. Introduction....135

5.3. Materials and Methods 136

5.3.1. Study participants 136

5.3.2. Discovery Mass Spectrometry Analysis 141

5.3.2.1. Data processing and statistical analysis for discovery lipidomic profiling 141

5.3.2.2. Lipid putative annotation 142

5.3.2.3. Selection of biomarker candidates for targeted validation 143

5.3.3. Targeted mass spectrometry analysis....144
5.3.3.1. Sample preparation....144
5.3.3.2. LC-MS/MS: targeted method....144
5.3.3.3. Statistical Analysis for targeted validation of lipid markers....148
5.4. Results....148
5.4.1. Selection of biomarker candidates for targeted validation....150
5.4.2. Targeted analysis and validation of biomarker panel....150
5.5. Discussion....156
Chapter 6.... 163
Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia.... 163
6.1. Abstract....164
6.2. Introduction....165
6.3. Materials and Methods....167
6.3.1. Materials and reagents....167
6.3.2. Animals....167
6.3.3. RUPP procedure....167
6.3.4. L-Ergothioneine in vivo experimental protocol....168
6.3.5. Plasma sample preparation....169
6.3.6. UPLC-MS analysis....169
6.3.7. Data processing....170
6.3.8. Statistical analysis....171
6.3.9. Putative annotation of metabolites....171
6.4. Results....172
6.4.1. RUPP control vs Sham control....172
6.4.2. RUPP +ET vs RUPP control....206

6.4.3. RUPP+ET vs Sham+ET 206

6.5. Discussion 211

Chapter 7: 216

General Discussion 216

7.1. Summary of main findings 217

7.2. Clinical implications of metabolomics findings 219

7.3. Thesis strengths 220

7.4. Thesis limitations 221

7.5. Future perspectives 224

7.6. Concluding remarks 225

References 227

Appendix A 257

Appendix B 260

Appendix C 274

Appendix D 281

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism.

Author contributions:

All experimental procedures described in this thesis were carried out precisely, and without bias, to the highest standards. The following methods were performed solely by Aude-Claire Morillon, with the exception of:

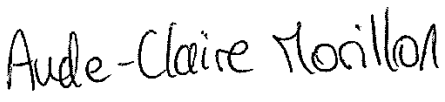
Chapter 2: Aude-Claire Morillon and Debora F Leite both carried out the systematic review and contributed equally.

Chapter 3: Debora F Leite assisted with Cork plasma samples preparation and analysis.

Chapter 3, 4 and 5: Gregoire Thomas performed the statistical analysis of the metabolomics data.

Chapter 5: The targeted analysis method was optimised by Kirsten Dowling. The samples for targeted analysis were prepared by both Aude-Claire Morillon and Kirsten Dowling.

Signed,



Aude-Claire Morillon

Acknowledgments

Firstly, I would like to sincerely thank my supervisors, Dr Fergus McCarthy, Dr Jane English, Dr Shirish Yakkundi, Prof. Louise Kenny, and Prof. Philip Baker. Thank you for giving me the opportunity to undertake this PhD and offering your guidance through the completion of this project.

Secondly, I would like to thank Science Foundation Ireland and Waters Corporation for funding this project, and the INFANT Centre for their constant support. I would like to thank the SCOPE consortium, all the pregnant women who participated in the SCOPE study, as well as all SCOPE research midwives and the whole SCOPE team.

Finally, I would like to thank my family, friends and colleagues. Without your never-ending support I would not have been able to finish my PhD. Especially Dr Alex Lee, Dr Indra San Lazaro-Campillo, Dr Marc O'Sullivan, and my mother, Mme Marie-Dominique Manceau. I am so grateful you were there when I need to vent, cry, and celebrate all the milestones along the way.

Thesis Abstract

Background

Clinical metabolomics is a growing field of research aiming to use metabolomic techniques to gain further knowledge into diseases, the use of biomarkers to predict their onset, or the effect of a potential therapeutic agent on the metabolome. Adverse pregnancy outcomes, such as small for gestation age (or fetal growth restriction), spontaneous preterm birth, and pre-eclampsia, lead to high maternal and fetal mortality and morbidity rates. However, despite research efforts to date, their pathophysiology remains poorly understood.

Aim

The aims of this thesis was to determine the accuracy of metabolomics to predict small for gestation age (SGA) babies, to explore the metabolic pathways involved in the pathophysiology of SGA and spontaneous preterm birth (sPTB), to identify potential predictive biomarkers of sPTB, and investigate the use of a potential therapeutic agent in an animal model of pre-eclampsia.

Methods

Firstly, a systematic review was undertaken to examine the predictive accuracy of metabolomics for the prediction of small for gestational age babies. The original search was conducted in February 2018 and the results are presented in **Chapter 2**.

Secondly, we investigated the metabolic pathways involved in the pathophysiology of small for gestation age (SGA) using untargeted ultra-performance liquid chromatography coupled to quadrupole time of flight mass spectrometry (UPLC-Q-TOF-MS). Plasma (Cork) and urine (Cork, Auckland) samples were collected at 20 weeks of gestation from pregnant women participating in the SCreening fOr Pregnancy Endpoints (SCOPE) study, an international study that recruited 5,628 nulliparous women, with a singleton low-risk pregnancy. Cases were women with

SGA (customised birthweight ≤ 10th centile) matched to controls who had uncomplicated pregnancies, according to age (±5 years), body mass index (BMI, ±3.5 kg/m2), and ethnicity. All samples were analysed in untargeted positive and negative ion modes, using UPLC-Q-TOF-MS. Data were processed, features were ranked based on p-values from empirical Bayes analysis adjusted for multiple testing, and significant features (adjusted p-values <0.05 were searched for identification (HMDB, LipidMaps)).

Thirdly, we aimed to decipher the lipidomics pathways involved in pathophysiology of spontaneous preterm birth (sPTB). Our analysis focused on plasma samples from SCOPE in Cork, collected at 20 weeks of gestation. Samples were profiled using semi-targeted liquid chromatography-mass spectrometry lipidomics, and lipids significantly altered between sPTB (n=16) and Control (n=32) groups were identified using empirical Bayes testing, adjusting for multiple comparisons. Significantly altered lipids (adj. p-values <0.05) were database searched for identifications (HMDB, LipidMaps).

Fourthly, in **Chapter 5**, we performed a discovery lipidomics experiment to determine potential biomarkers of sPTB, in plasma samples taken at 15 weeks of gestation in women who participated in SCOPE in Cork and Auckland. Selected participants were women who has sPTB before 34 weeks of gestation (n=16 from Cork, and n=23 from Auckland), matched to women who had an uncomplicated pregnancy (n=39) according to age (±5 years) and BMI (± 3 kg/m2). Lipidomics analysis was performed using UPLC-Q-TOF-MS. Statistical analysis using empirical Bayes, adjusted for multiple testing was used to create a list of potential biomarkers. Five potential biomarkers were selected for validation based on statistical analysis, and their identification was validated using standard mix and UPLC coupled to triple quadrupole mass spectrometer (TQ-MS) analyses. Their prediction potential was tested using samples taken at 15 and 20 weeks of gestation from women from SCOPE Cork who had sPTB before 37 weeks of gestation (n=54) matched to women who had an uncomplicated pregnancy (controls, n=108). In addition, plasma collected at time of delivery (ToD) was also analysed for six cases and their 12 matching controls. Cases were matched to controls according to age (±5 years) and BMI (± 3 kg/m2). Samples

were analysed using UPLC-TQ-MS, and statistical analysis was performed using independent T tests on normalised data. In addition, independent T tests were performed to determine if the levels of each target were significantly different between cases and controls at each time point (15 or 20 weeks). We defined significance as p-value <0.05.

Finally, in **chapter 6** we performed metabolomics analysis of plasma from experiments examining L-Ergothioneine treatment in the Reduced Uterine Perfusion Pressure (RUPP) rat model of pre-eclampsia. The effect of L-Ergothioneine (ET) treatment was explored using in vivo treatment in rats: Sham control (SC, n=5), RUPP control (RC, n=5), Sham + ET (ST, n=5), RUPP + ET (RT, n=5). Metabolic profiles of plasma samples were obtained using UPLC-Q-TOF-MS, and statistical analysis of the data was performed on normalised data, using independent T tests adjusted with false discovery rate (FDR) to compare RC to SC, RT to RC and RT to ST. Metabolites significantly altered (FDR <0.05) were putatively identified through database search (HMDB).

Results

The systematic review presented in **Chapter 2** examining the predictive accuracy of metabolomics for small for gestational age babies showed that to date no combination of metabolites are able to predict small for gestational age accurately. However, the review revealed the potential of investigating lipids pathways, their involvement in the pathophysiology of small for gestational age, and their high predictive potential.

The metabolomic studies performed on urine samples and reported in **Chapter 3**, showed lower levels of 4 metabolites of interest (sulfolithocholic acid, estriol-16-Glucuronide, Neuromedin N (1-4), and 4-Hydroxybenzaldehyde) in Cork were associated with SGA at 20 weeks of gestation, but not in Auckland samples. These urinary metabolites are associated with detoxification, nutrient transport and absorption pathways. The lipidomics analysis performed on plasma samples showed that higher levels of several glycerophospholipids (3 phosphatidylethanolamines, 5

phosphatidylserines, 3 phosphatidylcholines, 1 lyso phosphatidylcholine, 1 phosphatidylglycerophosphate, 1 lyso phosphatidylglycerophosphate, 2 phosphatidylinositols, 2 phosphatidylglycerophosphates, and 3 phosphatidylglycerols) in at 20 weeks of gestation were associated with the development of SGA in the Cork participants of the SCOPE pregnancy cohort.

Chapter 4 demonstrated that twenty-six lipids showed lower levels in sPTB compared to controls (adjusted p <0.05), including 20 glycerophospholipids (12 phosphatidylcholines, 7 phosphatidylethanolamines, 1 phosphatidylinositol) and 6 sphingolipids (2 ceramides and 4 sphingomyelines). In addition, a diaglyceride, DG (34:4), was detected in higher levels in sPTB compared to controls.

In **Chapter 4**, we reported that reduced levels of phospholipids (glycerophospholipids and sphingolipids) are associated with the pathophysiology of sPTB. In the UPLC-Q-TOF-MS discovery phase of the study presented in Chapter 5, a list of 120 potential lipid biomarkers were reported. Most were tentatively identified as glycerophospholipids and detected in lower levels in sPTB. From this list of features, 5 potential biomarkers predictive of sPTB were selected and used in a targeted UPLC-TQ-MS analysis. The results obtained showed that two of the targets showed significant differences between cases and controls and over time (between 15 and 20 weeks of gestation), PC (15:0/22:6) and TG (18:3/18:2/18:3).

In **Chapter 6**, using untargeted UPLC-Q-TOF-MS, we tested the effect of L-Ergothioneine (ET) as a potential therapeutic agent for the treatment of pre-eclampsia in the RUPP rat model. We reported significantly higher levels of L-palmitoylcarnitine, fatty acyl substrate involved in beta-oxidation in the mitochondria, in RUPP rats compared to Sham rats. When comparing plasma metabolic profiles of RUPP + ET rats to RUPP rats, we reported 10 metabolites associated with inflammation significantly altered (FDR <0.05, e.g. 20-COOH-leukotriene E4). Glutamylcysteine, a metabolite associated with oxidative stress, was detected at significantly higher levels (FDR <0.05) when comparing RUPP + ET rats to RUPP rats, and RUPP + ET rats to Sham + ET rats. These results show that the therapeutic properties of L-Ergothioneine might be related to mitochondrial function

preservation, by attenuating inflammatory response evident in pre-eclampsia in addition to increasing antioxidant levels.

Conclusions

Overall, these results show that glycerophospholipids appear to play a key role in the pathophysiology of SGA and sPTB, and dysregulated glycerophospholipids are potential makers of adverse pregnancy outcomes. Further research is needed to understand their precise associations, whether they are a cause or effect of SGA and sPTB, as well as to validate their potential as predictive biomarkers in independent pregnancy cohorts. In addition, we have shown that the use of L-Ergothioneine for the treatment of pre-eclampsia in the RUPP rat model reduces the oxidative stress induced by pre-eclampsia, via amino acid and glycerophospholipids metabolism pathways. Future work should focus on a testing L-Ergothioneine as a treatment for pre-eclampsia in a clinical trial.

This thesis has demonstrated the potential for metabolomics to help understand the pathophysiology of adverse pregnancy outcomes and has explored its use in assessing biological pathways, predictive biomarkers and potential therapeutic pharmacological interventions. To date results are limited with significant further validation required.

Contribution and List of Publications

This thesis contributed to further the field of metabolomics in pregnancy adverse outcomes research, by progressing the understanding of the pathophysiology of small for gestational age, and preterm birth. This work also identified two potential markers predictive of preterm birth and provided further evidence of the high potential of L-Ergothioneine as a treatment for pre-eclampsia.

Publications arising from this thesis:

<u>Published manuscripts</u>

- Leite Debora Farias Batista, **Morillon Aude-Claire**, Melo Júnior Elias F, Souza Renato T, Khashan Ali S, Baker Philip N, Kenny Louise C, Cecatti José Guilherme. Metabolomics for predicting fetal growth restriction: protocol for a systematic review and meta-analysis. ***BMJ Open*** 2018;8:e022743. doi: 10.1136/bmjopen-2018-022743
- Leite Debora Farias Batista, **Morillon Aude-Claire**, Melo Júnior EF, Souza Renato T, McCarthy, Fergus P, Khashan Ali S, Baker Philip N, Kenny Louise C, Cecatti José Guilherme. Examining the predictive accuracy of metabolomics for small-for-gestational-age babies: a systematic review. ***BMJ Open*** 2019;9:e031238. doi: 10.1136/bmjopen-2019-031238
- **Aude-Claire Morillon**, Shirish Yakkundi, Gregoire Thomas, Lee A Gethings, James I Langridge, Philip N Baker, Louise C Kenny, Jane A English, Fergus P McCarthy. Association between Phospholipid Metabolism in Plasma and Spontaneous Preterm Birth: A Discovery Lipidomic Analysis in the Cork Pregnancy Cohort. ***Metabolomics*** 2020 16(2): 19.
- **Aude-Claire Morillon**, Rachel D. Williamson, Philip N. Baker, Douglas B Kell, Louise C. Kenny, Jane A English, Fergus P. McCarthy, Cathal McCarthy. Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia. ***PLoS One*** 2020 15(3): e0230977.

- **Aude-Claire Morillon**, Debora F. B. Leite, Shirish Yakkundi, Lee A Gethings, Gregoire Thomas, Philip N. Baker, Louise C. Kenny, Jane A. English, Fergus P. McCarthy. Glycerophospholipid and Detoxification Pathways Associated with Small for Gestation Age Pathophysiology: Discovery Metabolomics Analysis in the SCOPE Cohort. ***Metabolomics*** *17(1), 5. doi:10.1007/s11306-020-01740-9.*

Manuscripts submitted for peer-review

- Kirsten G. Dowling, **Aude-Claire Morillon**, Gregoire Thomas, Daragh Boyle, Shirish Yakkundi, Billy Molloy, Lee A Gethings, James I Langridge, Louise C Kenny, Fergus P McCarthy, Jane A English. Altered plasma levels of triacylglycerol and phosphatidylcholine in preterm birth; a biomarker discovery and validation study in Cork SCOPE birth cohort. ***American Journal of Obstetrics and Gynecology.***

For more details about dissemination of my work, see **Appendix A**.

Abbreviations list

Abbreviations	Definition
AF	amniotic fluid
aHR	adjusted hazard ratio
aRR	adjusted risk ratio
CI	confidence interval
CPD-choline	Cytidine diphosphocholine
CRF	corticotropin-releasing factor
CRH	corticotropin-releasing hormone
DR	detection rate
IL	interleukin
ISSHP	international society for the study of hypertension in pregnancy
IUGR	intra-uterine growth restriction
FGR	fetal growth restriction
GC	gas chromatography
LC	liquid chromatography
LC-MS	liquid chromatography mass spectrometry
MAP	mean arterial pressure
MRM	multiple reaction monitoring
NAFLD	non-alcoholic fatty liver disease
NMR	nuclear magnetic resonance
OR	odds ratio
PCR	polymerase chain reaction
PEMT	phosphatidylethanolamine-N-methyltransferase
PlGF	placental growth factor
PPROM	preterm prelabour rupture of membranes
PTB	preterm birth
ROS	reactive oxygen species
SCOPE	screening for Pregnancy endpoints study
SGA	small for gestational age
sPTB	spontaneous preterm birth
SRM	selected reaction monitoring
sVEGFR-1	soluble receptor for vascular endothelial growth factor
sFlt-1	soluble fms-like tyrosine kinase-1
TCA cycle	tricarboxylic acid cycle
TLR	toll like receptor
TQ-MS	triple quadrupole mass spectrometry
UPLC-MS	ultra-performance liquid chromatography mass spectrometry
UtA-PI	uterine artery pulsatility index

Chapter 1:

Introduction

1.1. Small for gestational age

1.1.1. Definition and incidence

Small for Gestational Age (SGA) is defined as newborns whose birthweight is less than the 10th centile for their gestational age compared with fetal growth chart normalities (Kiserud et al. 2017). SGA can be further classified as moderate (birthweight in the 3rd-10th centile), or severe (birthweight less than the 3rd centile) (P. A. Lee et al. 2003). Fetal growth restriction (FGR), or intrauterine growth restriction (IUGR), is defined as impaired fetal growth compared with the normal growth potential for the fetus. An estimated 30 million neonates are affected by FGR every year (D. Sharma et al. 2016b). In spite of subtle differences, SGA can be considered a surrogate endpoint for FGR and is often used as a synonym (Rosenberg 2008). However, it is estimated that a third of SGA babies are actually constitutionally small, and are thus misdiagnosed (Lesley ME McCowan et al. 2005). Other definitions and synonyms of SGA exist and are listed in **Table 1.1**. To date there is no clear consensus as to the use of SGA for a proxy of FGR, but for the purpose of this thesis and for clarity, we have chosen to use the term SGA.

It is estimated that that SGA affect 10 % of pregnancies, and represents annually 13.7 million of term newborn babies (D. Sharma et al. 2016b). The exact incidence of SGA varies among countries, ethnicities, populations, and is higher with decreasing gestational age (D. Sharma et al. 2016b; M. De Onis et al. 1998). Asia is estimated to have the highest incidence of SGA, and accounts for nearly 75 % of all SGA, followed by Africa and Latin America (M. De Onis et al. 1998; D. Sharma et al. 2016b).

Table 1.1: Definitions and synonyms of small for gestational age.

Small for gestational age (SGA)	Birthweight below the 10th centile of population curves, or customised birthweight.
Intra uterine growth restriction (IUGR), or fetal growth restriction (FGR)	Impaired fetal growth compared with the normal growth potential for the fetus.
World Health Organisation	"Birthweight below the 10th percentile of the recommended gender-specific birthweight for gestational age reference curves." (Mercedes de Onis and Habicht 1996)
SGA, as defined in SCOPE study	Birthweight below the 10th centile of customised birthweight, according to mother's height, weight at 15 weeks visit, ethnicity, as well as sex and weight of the baby, and gestational age at delivery.
Individualised Birthweight Ratio (IBR)	"IBRs allow comparison of birthweights within a population whilst adjusting for factors known to influence fetal growth such as maternal age, parity, ethnicity and smoking status" (J. et al. 1995)
Customised birthweight	Birthweight adjusted for mother's height, weight ethnicity, sex and weight of baby, and gestational age at delivery of baby.
Population curves	Birthweight adjusted to a specific population parameters.

1.1.2. Impact on health

Babies born small for their gestational age have higher risks for short and long-term health. Indeed, several neonatal complications can be associated with SGA, such as perinatal asphyxia and neonatal adaptive complications (Rosenberg 2008), or stillbirth (Bukowski et al. 2014), thus increasing the mortality rate of these neonates (Simchen et al. 2000). In addition, premature SGA infants have higher risk of neonatal mortality compared to premature babies, and compared to term SGA neonates (Katz et al. 2013).

Infants born small for their gestational age are at higher risk of long-term morbidities (D. Sharma et al. 2016a), such as impaired neurodevelopment (Grantham-McGregor

1998) and growth (Chaudhari et al. 2008; Chaudhari et al. 2012), and early onset of adult disease in childhood (Calkins and Devaskar 2011), cardiovascular diseases, diabetes mellitus or hypercholesterolemia for example. Thus, SGA is a global public health concern. A more detailed understanding of the pathophysiology of SGA may help inform interventions to prevent this condition and help its clinical management.

1.1.3. Aetiology of Small for Gestational Age

A number of biological factors may lead to SGA (D. Sharma et al. 2016b). All factors negatively impacting the nutritional supply and respiratory functions of the fetus may lead to impaired fetal growth, and ultimately a small for gestational age infant. These fundamental biological functions are regulated on various levels, mainly genetic, placental, fetal and maternal.

1.1.3.1. Genetic factors

Genetic factors resulting in SGA may be maternal, placental or fetal (**Table 1.2**). Polymorphism in these genes may lead to errors in the synthesis of proteins and hormones they code for and can lead to SGA. Firstly, several maternal genes have been implicated, such as genes coding for leptin and endothelin 1 (ET-1) (Arslan et al. 2004; Nezar et al. 2009), visfatin (Malamitsi-Puchner et al. 2007), and thrombophilia genes (Rajewski and Skrzypczak 2006; Grandone et al. 2002). Secondly, placental genes have been linked with SGA pathophysiology, such as the gene coding for S110B protein (Florio et al. 2006) insulin-like growth factors 1 receptor (IGF-1R) mutations (Kawashima et al. 2012; A. C. Leal et al. 2013; A. d. C. Leal et al. 2011), N-terminal parathyroid hormone-related protein (Briana et al. 2007), and genetic deletion of IFG-1 and SHOX gene (Caliebe et al. 2012). Lastly, several placental genes have been implicated with SGA pathophysiology, such as placental growth factor (PlGF) (Vrachnis et al. 2013; S.-W. Li et al. 2014), Cullin genes (Gascoin-Lachambre et al. 2010), serpin peptidase inhibitor clade A member 3 (SERPINA3) genes (Chelbi et al. 2012), nuclear paraspeckle assembly transcript 1 (NEAT1) gene (Gremlich et al. 2014), insulin-like growth factors 1 and 2 (IGF1 and IGF2) and insulin-like growth factor binding protein 3 (IGFBP3) (Börzsönyi et al. 2011; M. H. Lee et al. 2010),

apoptosis Bcl-2 and Bax gene (Börzsönyi et al. 2013), homeobox genes (PADMA Murthi 2014; Padma Murthi et al. 2014; Padma Murthi et al. 2011; Padma Murthi et al. 2006), epidermal growth factor (EGF) (Rab et al. 2013), STOX1 gene (van Dijk and Oudejans 2011), and trophoblastic micro RNAs (miRNA-424 and miRNA-141) (Q. Tang et al. 2013; L. Huang et al. 2013).

Table 1.2: Genetic factors linked with small for gestational age infants.

Genes	Biological function
Maternal genes:	
- Leptin and endothelin 1 (ET-1)	Respectively, hormone regulating appetite, and vasoconstrictor.
- Visfatin	Adipocytokine associated with increased adipose depot in insulin resistance.
- Thrombophilia genes	Genes implicated in this blood clotting disorder.
Placental genes:	
- Placental growth factor (PlGF)	Critical role in angiogenesis and vasculogenesis of the placenta.
- Cullin genes	Involved in cell scaffolding and ubiquitin-proteasome system.
- Serpin peptidase inhibitor clade A member 3 (SERPINA3) genes	Serine protease inhibitor
- Nuclear paraspeckle assembly transcript 1 (NEAT1) gene	Retention of mRNAs to the nucleus.
- Insulin-like growth factors 1 and 2 (IGF1 and IGF2) and insulin-like growth factor binding protein 3 (IGFBP3)	Respectively, important role in growth, and involved in transport of IGF factors in the bloodstream.
- Apoptosis Bcl-2 and Bax gene	Respectively, anti- and pro- apoptotic genes.
- Homeobox genes	Important in development and preservation of lymphatic and vascular systems.
- Epidermal growth factor (EGF)	Promotes cell growth and differentiation.
- STOX1 gene	Mediator of trophoblast dysfunction.
- Trophoblastic micro RNAs (miRNAs), miRNA-424 and miRNA-141	Non-coding RNA regulating post-transcriptional gene expression.
Fetal genes:	
- S110B protein	Regulation of cell-cycle progression and differentiation, and other cellular processes.
- Insulin-like growth factors 1 receptor (IGF-1R) mutations	Tyrosine kinase transmembrane receptor involved in growth.
- N-terminal parathyroid hormone-related protein	Involved in cell growth and differentiation.
- Genetic deletion of Insulin-like growth factors 1 (IFG-1) and SHOX gene	Involved in growth.

1.1.3.2. Maternal factors

Maternal factors associated with SGA are listed in **Table 1.3**. Maternal diseases with negative impact on blood circulation will lead to an impaired uteroplacental blood flow, which in turn will result in SGA. Such maternal diseases include diabetes associated with vasculopathy, systemic lupus erythematosus, hypertensive disorders (gestational and non-gestational), sickle cell disease, chronic renal disease, and antiphospholipid syndrome. Ethnicity and maternal age, especially when extreme (less than 16 and over 35 years) have been associated with SGA (Fang et al. 1999; Strobino et al. 1995). Low body mass index (BMI) before and in late pregnancy has been associated with reduced birth weight (Nomura et al. 2012). SGA has also been associated with assisted reproductive technologies, both for singleton and multiple gestation pregnancies (Hendrix and Berghella 2008). Several maternal parasite infestations and infections have been associated with SGA, such as urinary tract infections, bacterial vaginosis, tuberculosis, TORCH, and malaria (Walker et al. 2014). Living in high altitude can lead to SGA; because of the reduced blood volume, the placenta does not receive enough oxygen (Krampl 2002). A low socioeconomic status has been shown to impact the nutrition of the mother and lead to poor nutrition, causing in turn SGA (Wilcox et al. 1995). Low socioeconomic status has also been linked with substance abuse that can lead to SGA, such as moderate alcohol intake (Lundsberg et al. 1997; Yang et al. 2001) and illicit drugs such as marijuana or cocaine (Zuckerman et al. 1989).

Table 1.3: Maternal factors linked with small for gestational age infants.

Maternal factors
Low body mass index
Extreme maternal age (younger than 16 or older than 35 years old)
Maternal medication
Maternal substance abuse
Socioeconomic status
Altitude
Ethnicity
Poor weight gain
Severe maternal starvation during pregnancy
Parity
Previous delivery of an SGA baby
Interpregnancy interval
Failure to obtain normal medical care during pregnancy
Hematologic and immunologic disorders
Maternal infections and parasite infestations
Pathological conditions in pregnancy such as pre-eclampsia and diabetes associated with vasculopathy
Maternal medical disorders (collagen vascular disease, nephropathy)
Assisted reproductive technologies

1.1.3.3. Fetal factors

A variety of fetal factors has been associated with SGA including both genetic and environmental factors (**Table 1.4**). Several metabolic disorders have been linked with SGA, such as Donohue syndrome, hypophosphatasia, galactosamia, agenesis of pancreas, generalised gangliosidosis type 1, fetal phenylketonuria, congenital absence of islet of Langerhans, I-cell disease and transient neonatal diabetes mellitus (Priyadarshi et al. 2015). About 3 % of SGA cases occur in multiple gestations, especially in monochorionic twins or a greater order of gestation, due to "uterine crowding". In such cases, one fetus will have a decreased growth rate from 28 weeks of gestation (Lewi et al. 2008; Fieni and Gramellini 2004). Congenital infections have been linked with about 5 % of SGA cases. In developed countries, due to vaccination policies in place, the rate of rubella cases has decreased, but cytomegalovirus and toxoplasmosis are still common. In developing countries, common congenital infections are syphilis, rubella, congenital HIV infection and malaria (Walker et al. 2014). About 1-2 % of SGA cases are associated with severe congenital anomalies, such as abdominal wall defects, for example omphalocele and gastroschisis,

congenital diaphragmatic hernia (Balayla and Abenhaim 2014), trachea-oesophageal fistula, congenital heart disease (Wallenstein et al. 2012) and neural tube defects including anencephaly. A recent population based study showed that among SGA babies, there is a malformation incidence of 22.3 % (Khoury et al. 1988). Several genetic syndromes have been shown to cause SGA, such as Brachmann de Lange syndrome, Cornelia de Lange syndrome, Bloom syndrome, Rubenstein-Taybi syndrome, Seckel syndrome, Fanconi syndrome, Russel-Silver syndrome, Roberts syndrome, Dubowitz syndrome, De Sanctis-Cacchione syndrome, Mulibrey Nanism syndrome, and Johanson-Blizzard syndrome (Hendrix and Berghella 2008). About 7-19 % of SGA babies have a chromosomal abnormality. The most common ones are trisomy 21 (Down syndrome), trisomy 18 (Edwards syndrome), and trisomy 13 (Patau syndrome) (Snijders et al. 1993; Anandakumar et al. 1996), but other chromosomal abnormalities are ring chromosomes, autosomal deletions, and uniparental disomy (Soong et al. 2009).

Table 1.4: Fetal factors linked with small for gestational age infants.

Fetal factors
Metabolic disorder
Multiple gestation
Congenital infection
Severe congenital anomalies
Genetic syndromes
Chromosomal abnormalities

1.1.3.4. Placental factors

The placenta is a barrier allowing exchange between fetal and maternal circulatory systems; the placenta serves as filter and has several physiological functions that will later be performed by the lungs, kidneys, gastrointestinal tract and endocrine glands (S. Zhang et al. 2015; Jansson and Powell 2007). In this context, it is clear that placental insufficiency, and any placental factors which have negative impact on blood, oxygen and nutrient supplies to the fetus (**Table 1.5**), can lead to fetal impaired growth (Dessì et al. 2015; S. Zhang et al. 2015).

In addition, chromosomal anomalies have been shown to impact fetal growth and has been associated with SGA, such as placental trisomy (Robinson et al. 2010), and confined placental mosaicism (Wilkins-Haug et al. 2006; Stipoljev et al. 2001). SGA has also been linked to placental infections, such as placental malaria (Conroy et al. 2013). SGA babies have placentae 24 % smaller than that of appropriate for gestational age (AGA) babies (Heinonen et al. 2001). This difference in placental weight leads to a decrease in villous surface area, as well as fetal capillary surface area, thus mother to fetus exchange of blood, nutrients and oxygen are reduced (Yu 1992). Such a decrease may contribute to impaired fetal growth. Abnormal placental vascularisation is also associated with SGA, some abnormalities, such as velamentous cord insertion, single umbilical artery, true knot, placental abruption and placental infarct, reduce nutrient transfer to the fetus (Bjoro 1983; Bjøro 1981; Djakovic et al. 2007).

Table 1.5: Placental factors linked with small for gestational age infants.

Placental factors
Aneuploidy disorders: - Placental trisomy - Confined placental mosaicism (CPM)
Abnormal placental vascularisation: - Abnormal uteroplacental vasculature - Avascular villi - Decidual or spiral artery arteritis - Thrombophilia-related uteroplacental pathology - Velamentous umbilical cord insertion - Abruptio placenta - Single umbilical artery - Syncytial knots - Multiple infarctions
Chronic inflammatory lesions
Placental hemangioma
Infectious villitis
Chronic villitis of unknown aetiology (CVUE)
Placental dysfunction (pre-eclampsia, pregnancy induced hypertension)
Partial molar pregnancy
Multiple gestation
Placental malaria

1.1.4. Predictors of Small for Gestational Age

Current clinical diagnosis of pregnancies complicated by SGA, and thus of SGA infants, is by ultrasound, but to date no method is able to predict SGA accurately. Indeed, in spite of numerous studies to identify biomarkers of SGA, there are currently none robust enough to be used in clinical setting (Conde-Agudelo A et al. 2013).

Metabolomics, as a moderately invasive approach, allows the detection of maternal metabolites during pregnancy to facilitate early intervention. In this context, a number of studies have been conducted, but to date none of the studies have identified robust metabolomics biomarkers for the prediction of SGA (Conde-Agudelo A et al. 2013). Nonetheless, these studies have identified various pathways that are altered in cases of SGA compared to healthy uncomplicated pregnancies, such as metabolites involved with glucose and fetal hypoglycaemia (Dessì et al.

2012), or TCA cycle intermediates, such as aspartate, alanine, glutamine and leucine (Dessì et al. 2015). A systematic review examining the predictive accuracy of metabolomics for small for gestational age babies is presented in **Chapter 2**.

1.2. Preterm Birth

1.2.1. Definition and incidence

According to the World Health Organisation (WHO), births occurring before 37 completed weeks of gestation, or less than 259 days from the first date of a woman's last menstrual period (LMP), are defined as preterm birth (PTB) (WHO 1977). This pregnancy adverse outcome is not defined by specific symptoms or signs, but by the gestational age at delivery occurring under 37 weeks (Kramer et al. 2012). Several phenotypic classifications of preterm birth exist (**Table 1.6**), but the most common include clinical presentation of PTB (induced, iatrogenic, spontaneous preterm labour, preterm prelabour rupture of membranes (PPROM)), pathological pathways (inflammation/infection, stress-induced, and vasculopathic pathways), and gestational age sub-groups (Kramer et al. 2012). The latter usually differentiates extreme prematurity (<28 weeks of gestation), severe prematurity (28-31 weeks of gestation), moderate prematurity (32-33 weeks of gestation), and near-term prematurity (34-36 weeks of gestation).

Preterm birth was estimated to have a global prevalence of 11.1 % (95 % CI: 9.1 %-13.4 %) in 2010 (Blencowe et al. 2012). The majority of cases of PTB occur in low and middle-income countries, with South Asian or sub-Saharan African countries making up 60 % (Blencowe et al. 2012). National estimates range from 5 % in some European countries to 18 % in some African countries (Blencowe et al. 2012). About 60-70 % of preterm births occur at 34-36 weeks of gestation, about 20 % at 32-33 weeks, about 15 % at 28-31 weeks, and about 5 % at less than 28 weeks (Robert L. Goldenberg et al. 2008). In addition, spontaneous preterm birth (sPTB) are estimated to occur in 31-50 % of all PTB, PPROM are estimated to occur in 6-40 %, and iatrogenic or induced PTB are estimated to occur in 30 % (Robert L. Goldenberg et al. 2008; Slattery and Morrison 2002).

Table 1.6: Phenotypic classifications of preterm birth.

Clinical presentation	- spontaneous preterm labour - preterm prelabour rupture of membranes (PPROM) - iatrogenic - induced
Pathological pathways	- inflammation/infection - stress-induced pathways - vasculopathic pathways
Gestational age	- <28 weeks: extreme prematurity - 28-31 weeks: severe prematurity - 32-33 weeks: moderate prematurity - 34-36 weeks: near term prematurity

1.2.2. Impact on health

Being born prematurely accounts for about 75 % of perinatal deaths (Slattery and Morrison 2002), among which infants delivered before 32 weeks of gestation account for 30-40 %. Recent global estimates show that in 2016, preterm birth was the main cause of death for children below 5 years of age; this represents approximately 35 % of newborn deaths, and 16 % deaths of children below 5 years of age (Hug L. et al. 2017).

Preterm birth can lead to short and long-term complications, such as neonatal respiratory conditions, neurological conditions, impaired neurodevelopmental outcomes, higher rates of hospitalisation, or learning difficulties (Platt 2014; Moreira et al. 2014). All of these complications can lead to considerable health system costs in the long-term, and cause great financial and psychological burden for the families of premature babies (Petrou et al. 2009; Singer et al. 1999).

1.2.3. Aetiology of Preterm Birth

The aetiology of PTB remains unclear, however, it is believed that this pregnancy complication is a syndrome with “multiple aetiologies” (Roberto Romero et al. 2006). Several pathways have been proposed to explain the onset of PTB, such as infection or inflammation, uteroplacental ischaemia, uterine overdistension, cervical

disorders, stress, and endocrine disorders (Roberto Romero et al. 2006; Robert L. Goldenberg et al. 2008). In addition, many fetal or maternal risk factors have been associated with PTB, as will be discussed further in **section 1.2.3.7** (Robert L. Goldenberg et al. 2008).

1.2.3.1. Infection/Inflammation

Based on microbiological studies, infection/inflammation is estimated to account for 25-40 % of PTB (Gonçalves et al. 2002; Romero et al. 1992b), with bacterial infections being the most common cause of infections associated with PTB (Roberto Romero et al. 2006). However, such numbers are probably underestimates because of the difficulties in detecting intrauterine infections with conventional culture techniques (Roberto Romero et al. 2006). The detection and identification of microorganisms in the amniotic cavity can be performed using more sensitive methods, such as detecting microbial DNA using polymerase chain reaction (PCR), or specific probes (Hitti et al. 1997).

The amniotic cavity is believed to be sterile; hence, detection of bacteria in the amniotic fluid (AF) is abnormal. Women who present with microbial invasion of the amniotic cavity are at higher risks of PTB, sPTB, developing clinical chorioamnionitis, and of having adverse perinatal outcome than women with PPROM or PTB but with sterile AF (Romero et al. 2002). The frequency of such microbial infections is inversely proportional to the gestational age at delivery for both sPTB and PPROM (Watts et al. 1992; Andrews et al. 1995).

Microorganisms can also colonise the decidua or the space between the amnion and the chorion, and not be detected in the AF cultures or PCR analysis (Andrews et al. 1995). Microorganisms can colonise the amniotic cavity and fetus following four pathways: (1) ascent from the vagina and the cervix; (2) haematogenous dissemination through the placenta (transplacental infection); (3) retrograde seeding from the peritoneal cavity through the Fallopian tube; and (4) accidental introduction at the time of invasive procedures (e.g. amniocentesis) (Romero et al. 2001). The most common pathway is the ascending route, and the hypothesis is that at

approximately 20 weeks of gestation when the membranes become tightly applied to the decidua, women who are colonised by microorganisms will become symptomatic and eventually progress to early PTB (Robert L Goldenberg et al. 2000).

Chronic intra-amniotic inflammation has been observed and is frequently associated with microbiological infection of the AF, as illustrated by the high concentrations of Interleukine-6 (IL-6) (Yoon et al. 1995; Roberto Romero et al. 1993). In addition, host defence against infection and modulation of the immune response in the genital tract of women is believed to be modulated by toll like receptor 4 (TLR-4) (Fazeli et al. 2005). Furthermore, *in vitro* studies have shown that, in the first trimester, TLR-4 ligation in trophoblast cells promotes the production of cytokines, and TLR-2 ligation induces apoptosis (Abrahams et al. 2004). Such evidence shows that through TLR-2, pathogens may directly promote cell death (Abrahams et al. 2004); this is observed in several pregnancy adverse outcomes including SGA (P Murthi et al. 2005; Huppertz et al. 2005), miscarriage (Vadillo et al. 2000), and pre-eclampsia (Huppertz et al. 2005). Additional evidence suggest that the innate immune system plays a role in parturition. Indeed, increased mRNA and protein expression of TLR-2 and TLR-4 in the chorioamniotic membranes are associated with spontaneous term labour and sPTB with histological chorioamnionitis (with or without ruptured membranes) (Y. M. Kim et al. 2004).

Fetal infection is the most serious and advanced stage of ascending intrauterine infections, and is associated with intra-amniotic inflammation and fetal inflammatory response (Romero et al. 1998). Fetal inflammatory response has been linked with PTB, and fetal injury and long-term complications, such as cerebral palsy, and lung disease (Yoon et al. 2000; Yoon et al. 1999).

1.2.3.2. Uteroplacental ischemia

Women who present with spontaneous preterm birth (sPTB) can be classified in two categories: (1) women with inflammatory lesions of the membranes and placenta, and (2) women without inflammatory lesions (R Romero et al. 1993b).

The main pathological features observed in the placenta of women presenting no evidence of inflammatory lesions are maternal and fetal vascular lesions (R Romero et al. 1993b). The maternal vascular lesions observed in the placenta of women with sPTB are atherosis, thrombosis of the spiral arteries, failure of physiological transformation of the myometrial segment of the spiral arteries, and a combination of the listed lesions (Roberto Romero et al. 2006). Fetal arterial thrombosis, and a decrease in the number of arteriole in the villi may be included in the fetal lesions (Roberto Romero et al. 2006).

Uteroplacental ischaemia, following maternal vascular lesions, is believed to be a possible cause of PTB. The exact mechanism leading to the onset of PTB after uteroplacental ischaemia is not yet understood. However, there is evidence that when the severity of uteroplacental ischaemia is marked, decidual necrosis and haemorrhage can occur, releasing thrombin, which in turn may activate physiological parturition pathways (Roberto Romero et al. 2006).

Fetal vascular lesions have not been studied as extensively as maternal vascular lesions, but are believed to induce fetal compromise, and lead to PTB. However, more evidence is needed to confirm these hypotheses.

1.2.3.3. Uterine overdistension

Despite both fetus and placental growth, intra-amniotic pressure is kept at a relative constant throughout the gestation (Sideris and Nicolaides 1990; Fisk et al. 1992). This is due to progressive myometrial relaxation, which is caused by progesterone and endogenous myometrial relaxants (e.g. nitric oxide) (Sladek et al. 1999). However, stretching can lead to increased prostaglandin release (Kloeck and Jung 1973), expression of gap junction protein or connexin-43 (Ou et al. 1997), increased oxytocin receptor in pregnant and non-pregnant myometrium (Ou et al. 1998), and increased myometrial contractility (Laudański and Rocki 1975). During pregnancy, progesterone inhibits expression of proteins and genes associated with the contraction resulting from myometrial stretching (Ou et al. 1997). The effects of this stretching increases in late gestation, and reaches a maximum during labour, as a

results of lower levels of circulating progesterone (Ou et al. 1997; Chow and Lye 1994).

At 25-29 weeks of gestation, the chorioamniotic membranes are distended by 40 %, by 60 % at 30-34 weeks of gestation, and by 70 % at term (Millar et al. 2000). Based on *in vitro* models of amnion cells (Maehara et al. 1996; Kanayama and Fukamizu 1989; Nemeth et al. 2000b) and fetal membrane explant (Nemeth et al. 2000a) studies, there is evidence that mechanical forces leading to uterine overdistension could initiate the activation of mechanisms that culminate in rupture of the membranes (Roberto Romero et al. 2006).

Preterm birth has been associated with uterine contraction frequency (J. Iams et al. 2002; Nageotte et al. 1988; Moore et al. 1994). However, there are wide variations in these frequencies in normal pregnancy, and a large overlap between women who delivered at term and women deliver prematurely; uterine contraction frequency is thus a poor predictor of PTB (J. Iams et al. 2002).

1.2.3.4. Cervical disorders

Cervical insufficiency is a syndrome associated with several pathologies, including recurrent pregnancy loss in the mid-trimester, specific forms of preterm labour with bulging membranes without rupture of membranes or significant uterine contractions, and may lead to precipitous term labour (J. D. Iams et al. 1995). Cervical disorders may result from traumatic damage to the structural integrity of the cervix, surgical trauma, or congenital disorders (R Romero et al. 1993a). The predominant cause of cervical insufficiency is cervical ripening, but it can be caused by other pathological process, during the mid-trimester, such as infection (Roberto Romero et al. 2006). Indeed, in close to 50 % of women presenting with clinical signs of acute cervical insufficiency, intrauterine infection was detected (Romero et al. 1992a).

1.2.3.5. Maternal stress

Maternal stress has been associated with PTB (Wadhwa et al. 2001; Hobel 2004). The timing and nature of the stressful stimuli is varied, and ranges from heavy workload to anxiety (Mozurkewich et al. 2000; Copper et al. 1996). The stressful insult could occur in pre-conceptual period or during pregnancy (Hobel et al. 1999; Copper et al. 1996). The exact mechanism leading stress to cause PTB is not fully understood, but corticotropin-releasing factor (CRF) seems to play a key role. It is proposed that, through the production of prostaglandins and cortisol, increased levels of CRF can induce parturition (Petraglia et al. 1990; Jones and Challis 1989; Sandman et al. 2006).

1.2.3.6. Endocrine disorders

Progesterone is a key hormone for pregnancy maintenance (Mesiano 2001), and it is hypothesised that its inhibition is central to trigger labour in humans (Roberto Romero et al. 2006). Indeed, progesterone decreases the production of chemokines (e.g. IL-8) by the chorioamniotic membranes (believed to be crucial for decidual/membrane activation), inhibits cervical ripening, down-regulates gap junction formation, and promotes myometrial quiescence (Kelly et al. 1992; Stjernholm et al. 1996; Chwalisz 1994; Gorodeski et al. 1987). However, the exact mechanism by which progesterone is involved in PTB remains unclear (Roberto Romero et al. 2006). The reduction of progesterone levels can also result from an intrauterine infection (Fidel Jr et al. 1998).

Progesterone administration to women with a history of previous PTB decreased the rates of PTB in a randomised clinical trial (da Fonseca et al. 2003). However, other clinical studies have conflicting results, and the mechanisms by which progesterone administration may reduce PTB rates remains unclear.

1.2.3.7. Risk factors associated with Preterm Birth

Several maternal and fetal factors have been associated with an increased risk of PTB, including maternal demographic characteristics, pregnancy history, nutritional status, present pregnancy characteristics, psychological characteristics, infection, adverse behaviours, uterine contractions, and cervical length (Robert L Goldenberg et al. 2005; Robert L. Goldenberg et al. 2008).

Risk factors associated with PTB are listed in **Table 1.7**. Lower socioeconomic and education status, and single marital status have been associated with higher risk of preterm birth (L. K. Smith et al. 2007; Thompson et al. 2006; Brett et al. 1997). The exact reasons for such associations are not clear, however, possible explanations include poorer quality and quantity of antenatal care, poorer nutritional status, increased cigarette smoking (Cnattingius 2004), greater use of illicit drugs (e.g. cocaine, cannabis, heavy alcohol drinking) (Forray 2016; Kesmodel et al. 2000), increased rates of SGA babies, physically demanding work, and increased rates of adverse psychological factors (Slattery and Morrison 2002). Ethnic origin is associated with PTB. In the USA and in the UK, the PTB rate for black women are in the range of 16-18 % compared to 5-9 % for white women of comparable age (Robert L. Goldenberg et al. 1996; Fiscella 1996).

Higher risks of PTB are associated with nulliparous adolescent women, especially when they are younger than 18 years old, across all age/parity categories (G. C. S. Smith and Pell 2001; Kozuki et al. 2013). Risks of PTB are also increased with advanced maternal age (≥45 years old) (Carolan 2013). History of previous preterm birth is associated with higher risk of PTB in subsequent pregnancy, as are maternal nulliparity and male babies (Ferrero et al. 2016). Multiple pregnancies represent 12-27 % of all PTB (Iannucci et al. 1996). Almost 60 % of twins are born prematurely, and 40 % will have spontaneous labour or PPROM. Other twins will have elective preterm delivery because of fetal or maternal disorders. Compared to singletons and twins conceived spontaneously, singletons and twins conceived through assisted conception technologies are at higher risk of PTB (J.-B. Qin et al. 2017b; J.-B. Qin et al. 2017a).

Close temporal proximity (<6 months, and <12 months) to a previous pregnancy is associated with an increased risk of PTB (Wendt et al. 2012). Potential explanations include maternal nutritional depletion, cervical insufficiency, folate depletion, and ascending intrauterine infection (Conde-Agudelo et al. 2012). Nutritional intake during pregnancy can be measured by using indicators, such as body mass index (BMI). Low BMI (<19 kg/m^2) is associated with a higher risk of PTB (Hendler et al. 2005). Short cervical length is associated with increased risk of PTB, and the commonly used cut-off is less than 25 mm; the risk increases with shorter cervical lengths (Committee on Practice Bulletins 2012).

Other risk factors for PTB include infections, such as chorioamnionitis, HIV, bacterial vaginosis, urinary tract infection, *Chlamydia trachomatis* infection, hepatitis C, syphilis, and malaria (Galinsky et al. 2013; Xiao et al. 2015; Haahr et al. 2016; Cunnington et al. 2013; Q.-t. Huang et al. 2015; J. Qin et al. 2014; Desai et al. 2007). Several maternal psychological disorders are linked to an increased likelihood of PTB, such as bipolar disorder, pregnancy-related depression, stress, and anxiety (Bodén et al. 2012; Staneva et al. 2015).

In addition, fetal, placental and uterine conditions, such as fetal birth defects, placenta previa, placental abruption, polyhydramnios, leiomyoma, and uterine anomalies are associated with PTB (Purisch et al. 2008; Crane et al. 1999; Ananth et al. 1999; Many et al. 1995; Y.-H. Chen et al. 2009; Fox et al. 2014).

Numerous risk factors have been linked with PTB, however, about 60 % cases of PTB remain without clear biological explanation (Ferrero et al. 2016).

Table 1.7: Risk factors associated with preterm birth.

Risk factors
Previous preterm delivery
Pathological conditions in pregnancy (e.g. pre-eclampsia, pre-gestational diabetes, and chronic hypertension)
Nulliparity
Male baby sex
Lower socioeconomic status
Lower education status
Single marital status
Maternal substance abuse (e.g. cannabis, cocaine, heavy alcohol use)
Smoking
Ethnicity
Extreme maternal age (<18 years old or ≥45 years old)
Inter-pregnancy interval <6 months and <12 months
Low body mass index (<19 kg/m^2)
Short cervical length (<25 mm)
Assisted reproductive technologies
Maternal infections and subsequent intrauterine transmission
Fetal, placental, uterine disorders

1.2.4. Predictors of Preterm Birth

Clinical risk factors are not able to achieve robust prediction of PTB as standalone tests (Ferrero et al. 2016), and neither are biochemical markers. While the aetiology is poorly understood, several biochemical markers have been shown to predict PTB with moderate accuracy (Menon et al. 2011; Lucaroni et al. 2018). Such markers include fetal fibronectin, a glycoprotein associated with extracellular matrix degradation. Fetal fibronectin is produced by amniocytes and cytotrophoblasts, and is physiologically present in cervicovaginal fluids before 22 weeks of gestation. However, its detection between 24 and 34 weeks of gestation has been associated with higher risks of PTB (Honest et al. 2002). A recent systematic review and meta-

analysis of randomised clinical trials has shown that the use of fetal fibronectin testing as a predictive biomarker of PTB has no association with the prevention of PTB in singleton pregnancies, but incurred higher costs (Berghella and Saccone 2016). Indeed, Berghella and colleagues have included the data from six randomised clinical trials, and their meta-analysis showed that fetal fibronectin testing in women with treated preterm labour showed similar incidence of PTB (<37 weeks of gestation) as compared to the control group (20.7 % vs 29.2), with a relative risk of 0.72 (95 % CI 0.52-1.01).

Several biochemical markers associated with immune function and inflammation have been investigated as potential biomarkers of PTB, such IL-6 and IL-8. A recent study has shown that increased levels of IL-6 (area under the curve (AUC) of 0.738) and IL-8 (AUC of 0.635) in cervicovaginal fluids predict PTB within 7 days, and especially when IL-8 levels are associated with cervical length parameter (AUC of 0.862) with a specificity of 92.8 % (Jung et al. 2016). This model, however, was not sensitive enough (56.4 %) to be used in clinical tests.

Despite promising biochemical markers of PTB, none has proven robust enough to accurately predict preterm birth in a clinical setting (Robert L Goldenberg et al. 1996; Robert L Goldenberg et al. 2005; Menon et al. 2011; Lucaroni et al. 2018). However, high-dimensional biology approaches, such as metabolomics, have great potential for identifying early prediction biomarkers of PTB (R Romero et al. 2006; Oskovi Kaplan and Ozgu-Erdinc 2018). In addition, metabolomics techniques are ideal to help clarify the pathophysiology underlying the onset of PTB (**Chapters 4 and 5**).

1.3. Pre-eclampsia

1.3.1. Definition and incidence

Pre-eclampsia is a complex pregnancy hypertension disorder characterised by *de novo* hypertension in pregnancy after 20 weeks of gestation. The International Society for the Study of Hypertension in Pregnancy (ISSHP) determined that pre-eclampsia "is diagnosed by the presence of *de novo* hypertension after 20 weeks of gestation accompanied by proteinuria and/or evidence of maternal acute kidney injury, liver dysfunction, neurological features, hemolysis or thrombocytopenia, and/or fetal growth restriction" (M. A. Brown et al. 2018).

It is estimated that over 70,000 maternal deaths and over 500,000 fetal and neonatal deaths are caused by pre-eclampsia globally every year (M. A. Brown et al. 2018). This pregnancy complication is estimated to occur in 5 % of nulliparous pregnancies, and is estimated to affect 4 million women each year worldwide (Magee et al. 2014).

1.3.2. Impact on health

In spite of research efforts, and novel therapeutic agents targeting various processes of the pathophysiology of pre-eclampsia, no cure has been found, and to date, delivery of the placenta is the only cure (Mol et al. 2016; Everett et al. 2012). However, timing of the delivery is crucial to balance the risks for mother and child. Maternal risks are usually the main criteria for such a decision, but in cases associated with growth restricted babies, the risks to the child might necessitate the induction of delivery (Mol et al. 2016). Delivery before 37 weeks of gestation is associated with risks, as described in **section 1.2**.

In addition to short-term maternal risks, women who had pre-eclampsia have a higher risk of developing cardiovascular diseases later in life (P. Wu et al. 2017). The systematic review and meta-analysis performed by Wu *et al.* has shown that PE is associated with increased risk of future heart failure (adjusted risk ratio (aRR) 4.19; 95 % confidence interval (CI) 2.09-8.38), stroke (aRR 1.81; 95 % CI 1.29-2.55) and future coronary heart disease (aRR 2.50; 95 % CI 1.43-4.37), as well as death from cardiovascular disease (aRR 2.21; 95 % CI 1.83-2.66) or coronary heart disease (aRR

2.10; 95 % CI 1.25-3.51). The exact mechanisms are not clear, however, the authors stated that significant confounders that were not measured might have contributed to the associations they presented.

Another long-term pathology has been associated with pre-eclampsia: end stage kidney disease (ESKD) (Khashan et al. 2019). In their systematic review and meta-analysis, Khashan *et al.* showed that ESKD was associated with women who had pre-eclampsia (adjusted hazard ratio (aHR) 4.96; 95 % CI 3.89-6.32; p <0.001), pre-eclampsia in two pregnancies (aHR 7.13; 95 % CI 3.12-16.31; p <0.001), and preterm pre-eclampsia (aHR 9.19; 95 % CI 5.16-15.61; p <0.001). These hazard ratios were determined by adjusting for several cofounders, but the authors noted that they might not have been able to account for all possible cofounders.

These associations show that women who experienced pre-eclampsia could benefit from preventive strategies, awareness, and lifestyle changes. Indeed, a recent study showed that lifestyle interventions such as smoking cessation, dietary habits changes, and exercise were able to reduce the risk of cardiovascular disease following pre-eclampsia by 4-13 % (Berks et al. 2013). However, intervention studies are needed before prevention and screening policies can be implemented.

1.3.3. Aetiology of Pre-eclampsia

Various theories and models have been proposed over the years, however, the exact cause and aetiology of pre-eclampsia remain unclear. Pre-eclampsia was initially believed to develop according to two stages: placental dysfunction, followed by a maternal syndrome. This two-stage model was first proposed and revised by Chris Redman and colleagues (C. W. G. Redman and Sargent 2009). Since then there is a general consensus in the field that the aetiology of PE is more complex, and new theories have been presented to try and explain what could be the cause of its onset, and the existence of various phenotypes (C. W. G. Redman 2014; C. W. G. Redman and Sargent 2009; Roberts and Hubel 2009). Among these new theories, Redman and colleagues have described a six-stage model of pre-eclampsia (C. W. G. Redman

2014). An overview of the six-stage model of pre-eclampsia is presented in **Figure 1.1**.

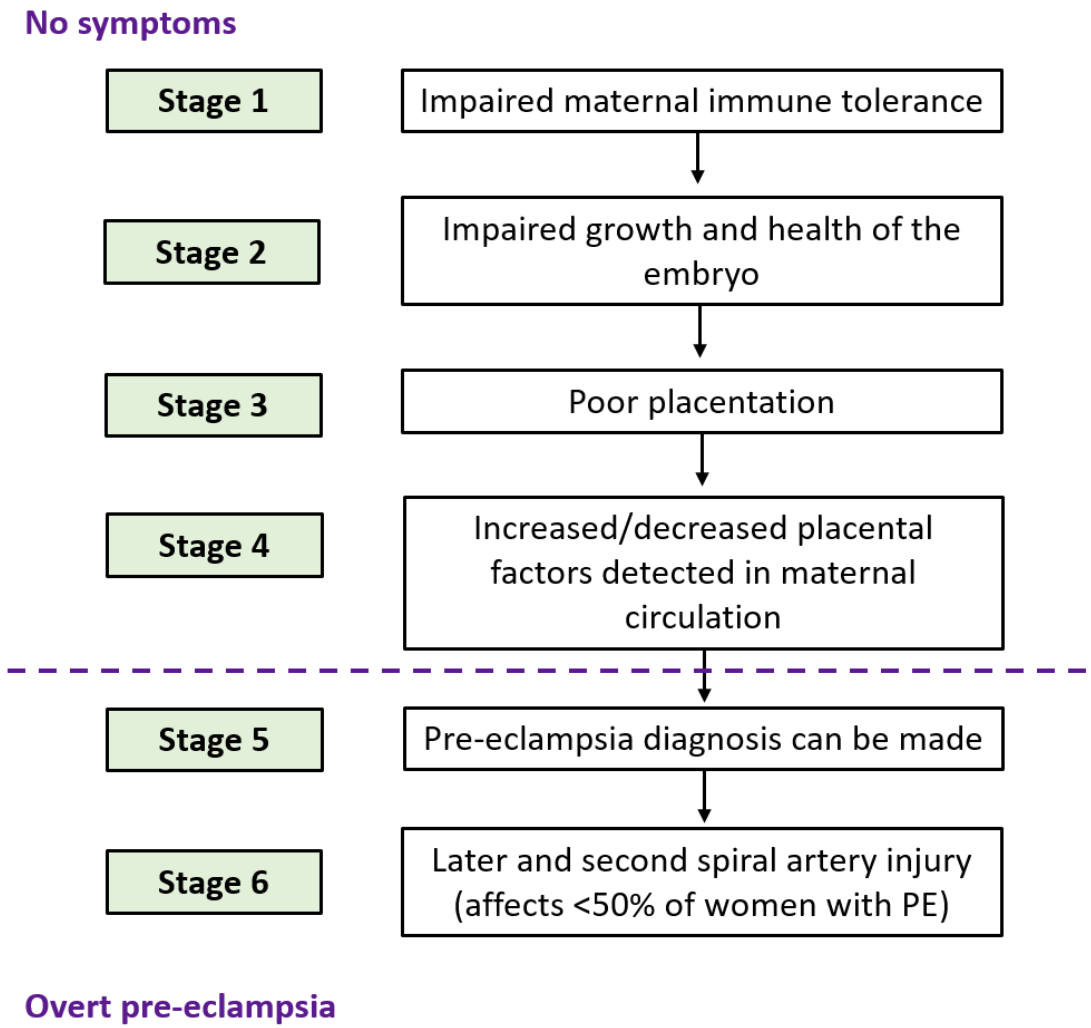


Figure 1.1: The six-stage model of pre-eclampsia, as described by Christopher Redman (C. W. G. Redman 2014).

In this model, the first stage of pre-eclampsia could stem from impaired immune tolerance (Saito et al. 2007), especially in regards to the preconceptual tolerisation of the mother towards the paternal semen. Short intervals between conception and first coitus are believed to cause a lack of tolerance of the mother towards the fetus' genetic foreignness, which in time can lead to impaired placentation (Christopher W.

G. Redman and Staff 2015; C. W. G. Redman and Sargent 2009; C. W. G. Redman 2014). However, another recent theory proposes that in some cases, the first stage of pre-eclampsia could be caused by microbial infection (Kell and Kenny 2016), which is believed to likely originate from the father through infected semen (Louise C Kenny and Kell 2018). Regardless of the cause, impaired immune tolerance is believed to adversely impact the growth and health of the embryo following implantation, which Redman considers to be stage 2 (C. W. G. Redman 2014). There is, however, lack of significant evidence to confirm this theory.

Stage 3 is believed to start with impaired placentation (C. W. G. Redman 2014), which starts after week 8 and lasts until week 18. Trophoblast invasion of the spiral arteries is associated with complete artery remodelling, and ensures sufficient blood perfusion to the uteroplacental unit (Christopher W. G. Redman and Staff 2015). However, when this remodelling fails, the blood flow is altered and over time leads to chorionic villi injuries, known as ischaemia-reperfusion injury (Burton et al. 2009).

Redman describes stage 4 as associated with various deficient or excessive placental factors being secreted form the injured chorionic villi over time, before clinical signs can be detected (C. W. G. Redman 2014). In addition, the poor placental perfusion has been shown to induce placental oxidative stress and increased production of reactive oxygen species (ROS) (Matsubara et al. 2015; Sánchez-Aranguren et al. 2014). ROS are highly reactive compounds known to damage DNA, RNA, and proteins, and an excessive production of ROS can overcome natural defences, which leads to oxidative stress (Birben et al. 2012). Recent studies have shown that oxidative stress mediated through the mitochondria, and mitochondrial dysfunction are implicated in the pathophysiology of pre-eclampsia (Williamson et al. 2017; Williamson et al. 2018; Vaka et al. 2018), and identified mitochondrial-targeted antioxidants as potential treatment for pre-eclampsia (C. McCarthy and Kenny 2016). L-Ergothioneine, a potential therapeutic agent targeting these cellular processes was investigated and the results are presented in **Chapter 6**.

L-Ergothioneine (ET) is a water soluble amino acid derivative of histidine that was first discovered in the ergot fungus by Charles Tanret in 1909 (Tanret C 1909; Borodina et al. 2020). ET exists as a tautomer and has both a thione and thiol forms

(Barger and Ewins 1911), the latter gives this molecule antioxidant properties (Franzoni et al. 2006; Servillo et al. 2015). In addition, this amino acid derivative has been detected in higher plants and animals, however, these organisms are not able to synthesise it (Borodina et al. 2020). High concentrations of ET have been detected in human tissues from dietary sources (Cheah and Halliwell 2012). This accumulation in human tissue is explained by the discovery of a solute carrier dedicated to the transport of ET: the solute carrier family 22, member 4 (SLC22A4) (Schomig 2005). The antioxidant properties of ET have been demonstrated in human subjects who were administered either 5 or 25 mg of ET once a day, for seven days (Cheah et al. 2017). The subjects showed reduced levels of blood markers of oxidative stress. Furthermore, several studies have shown evidence that ET accumulates in the mitochondrial fraction of hepatic cells in rats (Kawano et al. 1982; E. Smith et al. 2019).

Following placental ischaemia and oxidative stress, the syncytiotrophoblast sheds debris into the maternal circulation, which in turn initiates a systemic maternal inflammatory response. This inflammatory response induces the release of factors, such as the soluble receptor for vascular endothelial growth factor (sVEGFR-1, or sFlt-1) (C. W. G. Redman and Sargent 2009). The excess of circulating sFlt-1 binds to and inhibits vascular endothelial growth factor (VEGF), a factor necessary for the survival of the endothelium (Maynard et al. 2003), leading to systemic endothelial dysfunction. In addition, hypoxia was shown to induce an increased secretion of sFlt-1 from the placenta (Nevo et al. 2006).

Other pro-inflammatory factors derived from the syncytiotrophoblast could affect maternal systemic inflammatory response or angiogenesis, such as corticotropin releasing hormone (CRH) (Karalis et al. 1991), trophoblast-derived microparticles (Germain et al. 2007), placental growth factor (PlGF) (Torry et al. 2003), activin-A (K. L. Jones et al. 2007), and leptin (Matarese et al. 2005). All of these factors were detected in higher levels in PE (C. W. G. Redman and Sargent 2009). The maternal systemic inflammatory response and its network involved with the pathophysiology of PE include complement activation and leukocyte, coagulation function, the acute

phase response, hyperlipidaemia, and insulin resistance (C. W. G. Redman and Sargent 2009).

Redman states that when pre-eclampsia can be diagnosed, stage 5 has started, as opposed to stages 1 to 4 that happen without any clinical signs (C. W. G. Redman 2014). In addition, he believes that less than half of women diagnosed with PE will be affected by stage 6. This final stage is described as an additional spiral artery lesion, acute atherosis, which leads to an additional decrease of the uteroplacental perfusion (C. W. G. Redman 2014).

The aetiology of pre-eclampsia is still not fully understood; however, research efforts have been made to further establish the causes of this complex pregnancy complications. The reduced uterine perfusion pressure (RUPP) rat model has been well established to study the aetiology of pre-eclampsia (Vaka et al. 2018; Sedeek et al. 2008). In this rat model, pre-eclampsia in induced by a surgical procedure resulting in reduced blood flow into the uterus, which is performed to mimic the effects of ischaemia-reperfusion injury (Granger et al. 2006). In addition to the RUPP rats presenting pre-eclampsia phenotype, it has been shown that their pups are suffering from SGA (Walsh et al. 2009).

1.3.4. Risk factors associated with Pre-eclampsia

Several risk factors have been associated with pre-eclampsia (**Table 1.8**). Strong risk factors include previous pre-eclampsia or hypertension, diabetes (type 1 or type 2), autoimmune disorders (e.g. antiphospholipid syndrome, systemic lupus erythematosus), chronic kidney disease, and hypertension (National Institute for Health and Clinical Excellence (NICE) 2011; Bramham et al. 2014). Moderate risk factors include a family history of PE, nulliparity, multiple pregnancy, advanced maternal age (≥40 years old), high BMI (≥25 kg/m^2), inter-pregnancy interval greater than ten years, and polycystic ovarian syndrome (National Institute for Health and

Clinical Excellence (NICE) 2011; Roos et al. 2011). Other associated clinical risk factors include mean arterial pressure, mean uterine artery resistance, family history of coronary heart disease, maternal birth weight, vaginal bleeding for at least five days, cigarette smoking, alcohol use during the first trimester (North et al. 2011).

In addition, it has been shown that pre-eclampsia might be associated to a genetic predisposition (Williams and Pipkin 2011), including gene variants involved in inflammation, thrombophilia, the renin angiotensin system, and oxidative stress (Jebbink et al. 2012; Rana et al. 2014).

Table 1.8: Risk factors associated with pre-eclampsia.

Risk factors
Previous pre-eclampsia or hypertension
Pre-gestational diabetes (type 1 or type 2)
Autoimmune disorders (e.g. antiphospholipid syndrome, systemic lupus erythematosus)
Chronic kidney disease
Hypertension
Family history of pre-eclampsia
Nulliparity
Multiple pregnancy
Advanced maternal age (≥40 years old)
High body mass index (≥25 kg/m^2)
Inter-pregnancy interval >10 years
Mean arterial pressure
Mean uterine artery resistance
Family history of coronary heart disease
Maternal birth weight
Vaginal bleeding ≥5 days
Smoking
Alcohol use
Gene variants involved in inflammation, thrombophilia, the renin angiotensin system, and oxidative stress

1.3.5. Predictors of Pre-eclampsia

A systematic review reported that two biochemical markers involved in angiogenesis are associated with a higher risk of pre-eclampsia before 30 weeks of gestation, PlGF and sFlt-1 (Kleinrouweler et al. 2012). Kleinrouweler *et al.* estimated that, based on 15 included studies, the summary diagnostic odds ratio (OR) for lower serum levels of PlGF was 9.0 (95 % CI 5.6-14.5) for sensitivity and specificity of 0.74. From 19 included studies, Kleinrouweler *et al.* estimated that the summary diagnostic odds

ratio (OR) for higher levels of sFlt-1 was 6.6 (95 % CI 3.1-13.7) for sensitivity and specificity of 0.72. In addition, a recent review that assessed predictive and diagnostic tests for preterm pre-eclampsia found that lowered plasma levels of PlGF was a promising prediction tool, but still needed to be validated in clinical trials before implementation in routine care (Fergus P. McCarthy et al. 2018).

Using a combination of clinical risk factors, a prediction model of pre-eclampsia has been proposed, however, it only resulted in modest prediction of pre-eclampsia, with area under the receiver-operating characteristic curve (AUC) of 0.71 (North et al. 2011). Similarly, the full pre-eclampsia integrated estimate of risk (fullPIERS), a multivariate prediction model tool aiming to stratify patients within 48 hours of hospital admission was developed with AUC of 0.88 (95 % CI 0.84-0.92) (von Dadelszen et al. 2011). This prediction model is based on both clinical risk factors and biochemical markers, such as chest pain or dyspnoea, gestational age, serum creatinine concentration, aspartate transaminase concentration, and platelet count.

A model using a combination of proteomics biomarkers and clinical data to predict pre-eclampsia in low-risk women was recently published, using discovery and validation data sets (Myers et al. 2013). In this study, Myers *et al.* developed and validated 8 models that reached 50 % sensitivity, at 20 % positive predictive value and a prevalence of 5 % in their validation set, with similar performances. These 8 validated model combined blood pressure (mean arterial pressure at 20 weeks of gestation) to the following proteins: PlGF, insulin-like growth factor acid labile subunit, soluble endoglin, microtubule-associated protein RB/EB family member 1 or 3, disintegrin and metalloproteinase domain-containing protein 12, extracellular matrix protein 1, fructose-bisphosphate aldolase A, selenoprotein P, serine peptidase inhibitor Kunitz type 1, multimerin-2, and melanoma cell adhesion molecule. The authors note that this model needs validation in a larger cohort.

Another prediction model was developed using a combination of fourteen metabolites detected at 15 weeks of gestation, with AUC of 0.94, and AUC of 0.92 in independent samples (Louise C. Kenny et al. 2010). Furthermore, using a combination of clinical risk factors, eleven biomarkers, and uterine Doppler ultrasound at 20 weeks of gestation was shown to be a very good and accurate

prediction model for pre-eclampsia, with AUC of 0.90 (95 % CI 0.79-1.0) (Louise C Kenny et al. 2014).

All these prediction models seem promising, but still need further validation before they can be used in clinical practice.

1.4. Metabolomics

1.4.1. Metabolites, metabolome and metabolomics

Metabolomics is the study of all metabolites present in a biological sample at the time of sampling (Fiehn 2002). Metabolites are low molecular weight molecules produced by the cells and transformed by chemical reactions as part of cellular functions (Patti et al. 2012). All biochemical species are built from metabolite units, such as amino acids in proteins, nucleotides in DNA and RNA, and cell membranes (Dunn et al. 2011b). Metabolites are biochemical intermediates of cellular metabolic pathways. Their levels are regulated by enzymatic activities in response to cellular environmental changes, and they are the downstream result of gene expression (Goodacre et al. 2004). However, unlike genes and proteins, which are regulated by epigenetics and by post-translational modifications respectively (Dunn et al. 2011b; Villas-Boas et al. 2007), metabolites offer a direct reflection of the cellular activity on a functional level, and thus of the phenotype (Villas-Boas et al. 2007; Patti et al. 2012). Metabolites are a diverse class of compounds, their size ranges from 1 (proton) to more than 1500 Da (e.g. lipids and small peptides); their physicochemical properties include polarity, hydrophilicity and hydrophobicity, boiling point, molecular weight, acidity and basicity. Metabolites include mainly organic (amino acids, lipids, vitamins, carbohydrates, fatty acids), but also inorganic compounds, such as metals (Dunn et al. 2011b). In this context, the complete coverage of all present metabolites make the analysis technically challenging, and requires a combination of experimental strategies, including different analytical methods. Indeed, the most common metabolites are classified according to their chemical structure (e.g. lipids) or by their involvement in a specific pathway (e.g. glycolysis) (Dunn et al. 2011b).

The metabolome is defined as “the complete set of all metabolites formed by the cell in association with its metabolism” (Villas-Boas et al. 2007). The metabolites that compose it are part of metabolic pathways involved in catabolism (production of energy by breaking down of complex molecules) and in anabolism (synthesis of molecules from simple compounds to store energy) (Dunn et al. 2011b; Villas-Boas et al. 2007). Several metabolomes exist depending on the origin of the metabolites,

such as endogenous compounds synthetized and utilised by the organism, and exogenous compounds originating from the environment (e.g. nutrients, drugs, or pollutants) and metabolised by the organism, making up the endometabolome and exometabolome respectively (Dunn et al. 2011b; Villas-Boas et al. 2007). The metabolome can also be described according to the sample type analysed, and in the case of biological fluids such as urine or plasma, the metabolome is composed of metabolites originating from several organs and tissues, it can be used as a snapshot of the whole biological system (Dunn et al. 2011b). Metabolites are part of several interconnected metabolomes and metabolic pathways, and in fact, metabolic models have demonstrated that less than 30 % of metabolites are only involved in two reactions (Villas-Boas et al. 2007). With this level of diversity and interconnection, it is important to use informatics resources to gain more information about the nature of the metabolites studied and which reactions they are involved in, such as the Human Metabolome database (HMDB) (David S. Wishart et al. 2018), the Small Molecule Pathway Database (SMPDB) (Frolkis et al. 2010; Jewison et al. 2014), and the Kyoto Encyclopedia of Genes and Genomes (KEEG) Pathway maps (KEGG).

Metabolomics is the study of the metabolome and aims to detect and quantify metabolites present in the biological system studied, and allows for the sensitive measure of a specific phenotype, which arises in response to genetic or environmental variations (Dunn et al. 2011b). The metabolome is viewed as downstream information, flowing from the genome to the transcriptome, to the proteome, to the metabolome. It has been shown that, unlike the metabolite concentrations, metabolite flux (rate of synthesis or consumption) is only weakly impacted by enzyme concentration (Raamsdonk et al. 2001). Furthermore, the flux of metabolites can be measured in seconds, as opposed to minutes or hours for turnover rate in transcriptome and proteome, which makes the metabolome an ideal indicator of any environmental and phenotypical modifications (Dunn et al. 2011b).

Metabolomics can be applied to a wide range of scientific fields, such as microbiology, environment, or health, and a range of research questions, including better understand metabolic flux, or what is the effect of a specific stimuli on the

metabolome. The focus in this thesis is on clinical metabolomics (**section 1.4.4**), particularly in regards to biomarker discovery and the effects of a potential therapeutic treatment on the metabolome, using mass spectrometry techniques. Clinical metabolomics refers to the use of metabolomics experiments in the study of diseases, including pathophysiology, biomarkers discovery, and potential therapeutics treatments. Biomarkers are molecules found in biological fluids that relate to normal or abnormal process, or to a specific condition or disease.

Before expanding on the various clinical applications of metabolomics, it is necessary to present the analytical platforms, and the typical analytical workflow used in metabolomics.

1.4.2. Metabolomics analytical platforms

Several analytical platforms can be used in metabolomics, and they all present advantages and limitations, which need to be taken into consideration depending on the aim of the study. Mass spectrometry (MS) and Nuclear Magnetic Resonance (NMR) spectroscopy are the two techniques used the most frequently in metabolomics biomarker discovery, or untargeted studies (Dunn et al. 2011b), and are the two techniques this section will focus on.

1.4.2.1. Mass spectrometry

Mass spectrometry is an analytical technique capable of separating metabolites that have been ionised, either positively by addition of a proton or negatively by removal of a proton, and then separated on their mass to charge ratio (m/z). The separation and detection of the ionised metabolites is performed under high vacuum pressure to minimise the number of collisions between ions and molecules, which can negatively affect the mass accuracy, mass resolution and sensitivity of the mass spectrometer. Several types of mass spectrometers exist, but all are designed with four main components (**Figure 1.2**): the ion source, the mass analyser, the detector, and the data collection module (computer). From the sample, the metabolites are ionised in the source, then separated according to their m/z in the analyser, and because of their low molecular weight, most metabolites are singularly charged.

Finally, ionised metabolites reach the detector and the data are collected in the form of spectrum of m/z values corresponding to their intensity. Mass spectrometers can be set to scan any mass range of interest; in metabolomics this range is typically from 50 to 1500 Da.

In clinical metabolomics, MS is typically used on gaseous or liquid biological samples, for which soft ionisation techniques that preserve the integrity of the molecules are preferred. These include high vacuum pressure ionisation with matrix-assisted laser desorption ionisation (MALDI), and atmospheric pressure with electrospray ionisation (ESI).

Discovery metabolomics requires high mass accuracy and mass resolution mass spectrometry. Mass accuracy is defined as the error between the theoretical mass of a metabolite and the determined mass. Mass resolution is defined as the mass peak width, as calculated for Full Width Height Maximum (FWHM). With higher mass resolutions, peak widths are narrower, and the separation between metabolites of similar but not identical accurate mass is possible. Instruments with high accuracy and high resolution offer the ability to detect ions with similar accurate mass separately, and thus determine the accurate mass for metabolites putative identification. Such high specification instruments include Time-of-flight (TOF), quadrupole (Q), and hybrids (e.g. Q-TOF, triple quadrupole (TQ)), which also offer high sensitivity, fast scan acquisition rates.

The main advantages of MS are its ability to detect hundreds of thousands of metabolites in a sample when combined with chromatography, its high sensitivity, and its capability to obtain chemical identification. However, automated identification of features is still complex, and the identification of all detected features remains currently impossible.

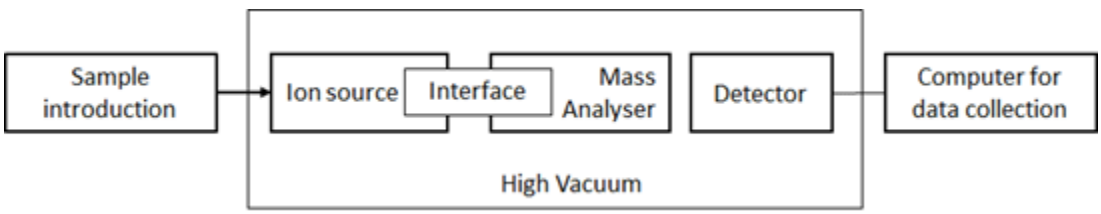


Figure 1.2: Schematic of the main components of a mass spectrometer.

1.4.2.2. Liquid chromatography mass spectrometry (LC-MS)

The development of atmospheric pressure electrospray ionisation (ESI) gave rise to the routine use of liquid chromatography mass spectrometry (LC-MS) in metabolomics studies. LC-MS is one of the most commonly used methods in clinical metabolomics, where the biological samples used are often complex biological matrix. Using LC allows for the separation of the metabolites present in these complex samples prior to their injection and analysis in the mass spectrometer. Indeed, LC offers separations of metabolites between a liquid mobile phase (containing the sample) and a solid (or liquid) stationary phase. The stationary phase is usually a LC column packed with particles on which the stationary phase is present. The column dimensions (length, diameter), stationary phase, mobile phase flow rate and temperature all have an impact on the chromatographic resolution and peak width. The introduction of ultra-performance liquid chromatography (UPLC) fifteen years ago, has led to narrower peak widths by using increased pressures and higher flow rates on smaller diameter column packing (Swartz 2005). UPLC can offer higher sensitivity than conventional LC, and chromatographic resolution equivalent to that of gas chromatography (GC), which typically offers peak width of less than 3 seconds. Using UPLC-MS, it is possible to detect thousands of features in a sample, and various types of columns can be used, the most common being reverse-phase C_8 or C_{18} columns. Reverse-phase separation is especially suited for relatively non-polar metabolites, including lipids, as it uses an elution starting with a high content in water and a gradient increase of the organic solvent (e.g. acetonitrile or methanol). Reverse-phase separation is, however, not best suited for polar metabolites, such as sugars and some amino acids, which tend to interact less with the solid phase column and eluate early when the elution gradient contains mostly water. A recent method has been developed, and continues to be optimised for the specific analysis of polar metabolites: hydrophilic interaction chromatography (HILIC). HILIC analyses use a hydrated silica column and an elution gradient starting with high organic content and increasing towards high aqueous content, which achieves the separation of more

polar metabolites and poor retention of non-polar lipids (Gika et al. 2008). The choice of column and gradient of elution thus depends on the research question.

Waters ACQUITY UPLC coupled to a Synapt G2-S quadrupole time-of-flight mass spectrometer (UPLC-Q-TOF-MS) and ACQUITY UPLC coupled to a Xevo triple quadrupole mass spectrometer (UPLC-TQ-MS) were the instruments available for the untargeted and targeted experiments presented in this thesis. These platforms will be further discussed in **sections 1.4.4.1** and **1.4.4.2**. These two analytical platforms can achieve high coverage of the metabolome and high mass accuracy and sensitivity, which are both necessary for untargeted and targeted clinical metabolomics experiments presented in this thesis.

1.4.2.3. Other analytical platforms frequently used in metabolomics

Gas chromatography mass spectrometry (GC-MS)

Gas chromatography (GC) has been used in combination with mass spectrometry for 60 years, and provides high chromatographic resolution (peak width below 3 seconds) and high sensitivity (detection limits in pmol or nmol range). GC-MS is mainly used in plant, microbial and clinical metabolomics applications. Chromatographic separation is obtained by using capillary columns with a stationary phase coated on the inner surface through which an inert carrier gas flows, and is placed in an oven (≤ 350 °C). Optimisation of the analytical conditions are obtained by increasing the column temperature, and varying the polarity of the stationary phase.

GC-MS operated by vaporising the sample into a gas prior to injecting into the capillary column, which means only the metabolites with a low boiling point will be vaporised. The majority of the endogenous metabolites have a low volatility, thus chemical derivatization is typically applied to increase the range of detectable metabolites by GC-MS techniques. The final chromatograms are complex and contain peaks from hundreds of metabolites, as well as numerous derivatization products.

Direct infusion mass spectrometry (DI-MS)

Chromatographic separation prior to MS detection is not always necessary depending on the samples type and the research question. Direct infusion mass spectrometry (DI-MS) is used on mass spectrometers fitted with an electrospray ionisation (ESI) source. Samples are directly injected into the mass spectrometer, and a single or averaged mass spectrum per sample is produced. The complexity of the metabolome requires that the mass spectrometer used in DI-MS can offer suitable separation of the majority of the metabolites detected, with high mass accuracy and resolution. High-throughput can be achieved with DI-MS, where up to one sample per minute can be analysed, however, as the full complex matrix sample is injected at once, the capacity for definitive metabolite identification is reduced and ion suppression increases.

Nuclear magnetic resonance spectroscopy (NMR)

For almost 30 years, NMR spectroscopy has been an invaluable tool for structural biologists and chemists, as well as metabolic profiling studies. A large number of metabolites can be detected simultaneously by NMR spectroscopy, since all cellular metabolites contain ^{1}H, as well as additional observable nuclei, such as ^{13}C or ^{31}P. Indeed, NMR techniques are based on the nuclear spin properties that certain nuclei possess, which means these nuclei can exist at different energy levels in a strong magnetic field (Goodacre et al. 2004). Observing these shift in energy levels generates invaluable information about chemical structure of the detected metabolites (Dunn et al. 2011b). This technique is quantitative, highly reproducible, and contrary to MS, its sensitivity is not dependent on polarity or the pK_a of the metabolites analysed. However, the sensitivity is low due to the small difference in energy between base energy and excited stated energy of the nuclei. This low sensitivity means that only metabolites present in high concentrations can be reliably quantified (Dunn et al. 2011b), for example when using a simple one-dimensional

sequence, only 20 to 30 metabolites are detected in serum or plasma samples, and 30 to 100 in urine samples.

NMR based metabolomics studies typically use samples in a solution state, such as biofluids (urine, plasma) or solubilised metabolites extracted from tissue samples. Sample preparation for NMR analysis are minimal and non-destructive, which makes multiple analysis of the same sample possible. In addition, the use of automated injections from 96 well plates and higher magnetic fields has enabled rapid high throughput (L. Zhang et al. 2016).

1.4.3. Metabolomics workflow

A generic metabolomics workflow (**Figure 1.3**) is followed when performing metabolomics experiments (M. Brown et al. 2005), and each step is optimised for the specific study question, which includes the sample preparation method, the type of instrumentation used, and bioinformatics approaches (Dunn et al. 2011b). As the work presented in this thesis only pertains to liquid chromatography coupled to MS experiments of biological samples, this section will focus on these specific techniques. The first step of untargeted analysis workflow (e.g. for biomarker discovery) is the analysis of samples, and the generation of the spectral data, which generates metabolic features information (e.g. m/z, retention time (RT)). These data are referred to as raw data, and are processed to ensure a reduction of the complexity and determine a robust set of features are detected across all samples (M. Brown et al. 2005; Dunn et al. 2011b). These features are then subjected to statistical analysis, and metabolite identification, after which biological interpretation of the metabolites of interest can be performed. In the case of targeted analysis, and after biological interpretation from untargeted experiments, metabolites of interest are selected for further analysis and validation in independent samples. The workflow applied to the data obtained from targeted analysis is very similar to that of the untargeted data.

The following section will present an overview of each of these steps, however, more details can be found in Methods **sections 3.3**, **4.3**, **5.3** and **6.3**.

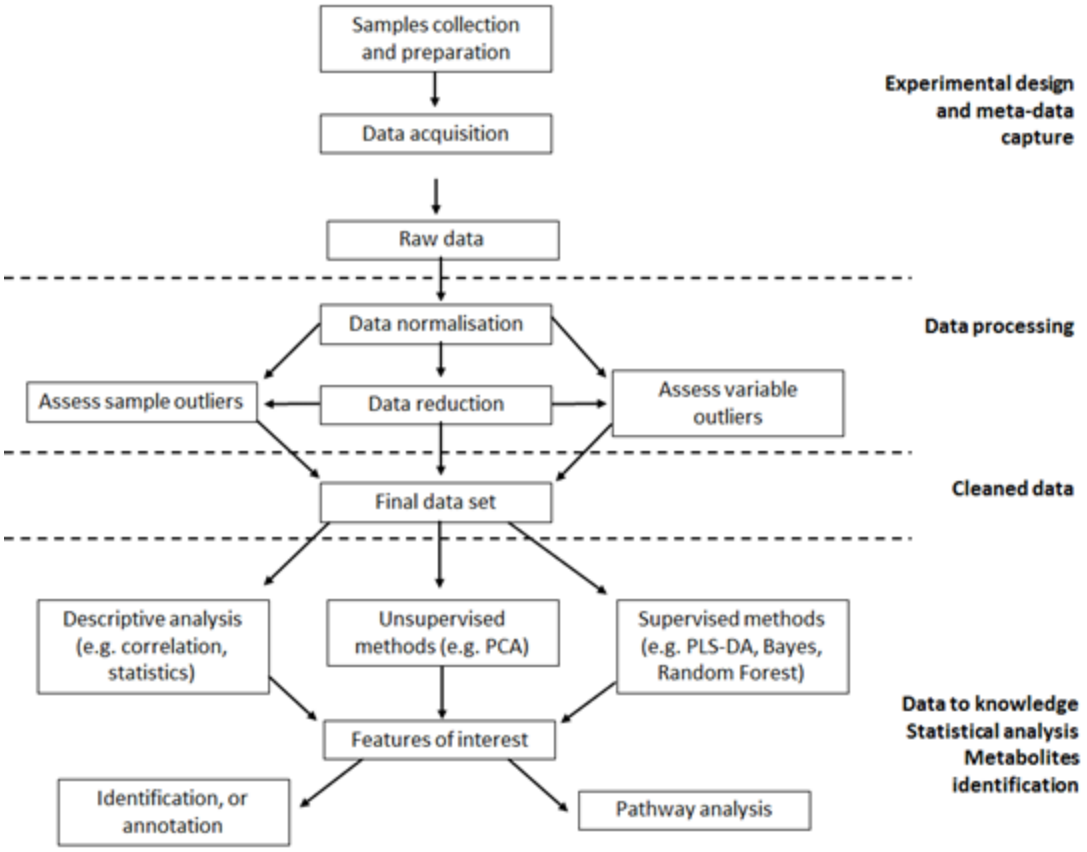


Figure 1.3 : Experimental workflow used in untargeted metabolomics. Modified from Brown *et al.* 2005 (M. Brown et al. 2005).

1.4.3.1. Sample collection, sample preparation and data acquisition

To ensure the samples analysed are representative of the metabolome at the time of collection, quenching of the metabolism, especially enzymatic activity, is paramount. Quenching of biological fluids is typically performed by decreasing their temperature, or adding organic solvent and/or applying heat in order to chemically inactivate enzymes. In order to analyse the metabolites present in a sample, they usually need to be separated from the matrix. Various extraction methods can be used depending on the type of sample, the experimental strategy (untargeted or targeted analysis), and the type of analytical platform used. In brief, urine samples collected from healthy subjects usually have low protein content, and samples

preparation is simple and consists mainly of dilution before analysis (Want et al. 2010). However, plasma samples contain high concentrations of proteins, and thus, deproteinisation needs to be performed prior to metabolites extraction (Dunn et al. 2011a).

Once metabolite extraction has been performed, samples are ready to be analysed and data is acquired. Data obtained with MS platforms generates complex data, called raw data (**section 1.4.3**).

1.4.3.2. Bioinformatics for metabolomics

Data processing

Because of the complexity of the raw data, processing steps are performed to ensure a reduction of its complexity, and the selection of robust features or metabolites in all samples (M. Brown et al. 2005). These steps might vary depending on the analytical platform used, but generally include spectral alignment, deconvolution, normalisation, and feature detection (or peak picking) (Dunn et al. 2011b). In the case of LC-MS, the processed data obtained are arranged as a matrix of feature intensity, ready for statistical analysis.

Pre-treatment of the data can be necessary, and consists of data normalisation, scaling, transformation (van den Berg et al. 2006), imputing missing values (Little and Rubin 2019), and detection and removal of outliers (Lindon et al. 2006). In addition, signal correction can be implemented to try and reduce bias in the data set, such as instrument drift. Once performed, the data is deemed “clean” and ready for statistical analysis.

Depending on the analytical platform used, various software and packages are available to perform the data processing steps. For the experiments presented in this thesis, the data was processed using the software package developed for Waters instruments, Progenesis QI (version 2.4), which uses accurate mass measurements, isotopic fit and fragmentation data to accurately process LC-MS data.

Statistical analysis

The type of statistical analysis implemented will depend on the hypothesis tested by the analytical experiment. In clinical metabolomics, such hypotheses are often questioning whether a significant difference will be detected between "healthy" and "diseased" subjects, or between subject "treated" or "not treated" with a pharmacological agent. Features that are determined to be significantly altered can then be further tested in targeted analyses, where validation of the findings will be performed. Each metabolite detected can be analysed in turn by univariate tests, such as Student's T test, ANOVA, or the non-parametric Kruskal-Wallis, or Mann-Whitney. Multivariate statistical methods can be applied to visualise the complex metabolomic dataset in a comprehensive manner, where all metabolites are simultaneously tested. Such methods are used to correlate biological variables with patterns in the metabolic features, and can be supervised or unsupervised. Unsupervised multivariate methods are used to recognised pattern in the data without knowledge about the phenotype or sample class. Principal component analysis (PCA) is the most common method used in metabolomics, and is used in this thesis. PCA helps visualise the data by using the linear transformation of the metabolites as a base to create an axis of linear uncorrelated variables (principal component) that explain maximal variance in the data (Jolliffe 1990). PCA score plots can help visualise global changes in the metabolome, and the majority of the variables responsible for the variances observed are usually described in the first few principal components (PC), especially in the first one (PC1). In addition, PCA score plots help visualise the technical quality of an analysis, by visualising the quality check control samples and potential outliers.

Supervised multivariate analyses are particularly interesting in clinical metabolomics, as these statistical methods correlate patterns of metabolites with the knowledge of phenotype. This enables the discrimination of biomarker signature (or combination of metabolites) as significantly associated with the studied phenotype, when single metabolites may not show significant association. The most commonly used multivariate supervised algorithm used in metabolomics is partial least squares discriminant analysis (PLS-DA) (Wold et al. 2001). It is, however, important to note

that model validations are necessary for these very powerful algorithms as they can find random associations easily.

Metabolite identification

The metabolomics standards initiative (MSI) has defined two types of identifications, definitive and putative, further classified into four levels (Fiehn et al. 2007; Salek et al. 2013; Sumner et al. 2007). These four levels are: identified metabolites (level 1), putatively annotated compounds (level 2), putatively characterised compound classes (level 3), and unknown compounds (level 4). In order to reach level 1 identification, researchers need to match two or more properties (e.g. accurate mass, retention time, isotope pattern, etc.) of the experimental data they acquired to an authentic chemical standard analysed in the same laboratory, using the same instrumentation and experimental conditions. To reach levels 2 and 3, annotation it is not required to use authentic standards. Most of the work presented in this thesis reach putative annotation level 2. The targeted study presented in **Chapter 5** reaches identification level 1 for the targeted metabolites.

Putative annotation is usually performed by matching experimental parameters (e.g. accurate mass, isotope pattern, etc.) to that of metabolites in specific databases. Several biological databases have been developed and curated over the years and keep expanding as more molecules and metabolites are detected and quantified. Such databases are the human metabolome database (HMDB) (David S. Wishart et al. 2018), lipid metabolites and pathway strategy database (Lipid Maps) (Cotter et al. 2007), or the Kyoto encyclopedia of genes and genome database (KEGG) (KEGG). In addition to chemical information, these databases contain information about the pathways each known metabolites are involved in, which is crucial information to provide biological context and meaning to the experimental data. With the development of metabolomics field, a recent database specialised in the pathways of small molecules was created, the small molecule pathway database (SMPDB) (Frolkis et al. 2010).

1.4.4. Workflow for clinical metabolomics

Two main mass spectrometry approaches are used in the investigation of clinical metabolomics: untargeted metabolic profiling (or biomarker discovery), and targeted analysis of metabolites (biomarker validation) (**Figure 1.4**).

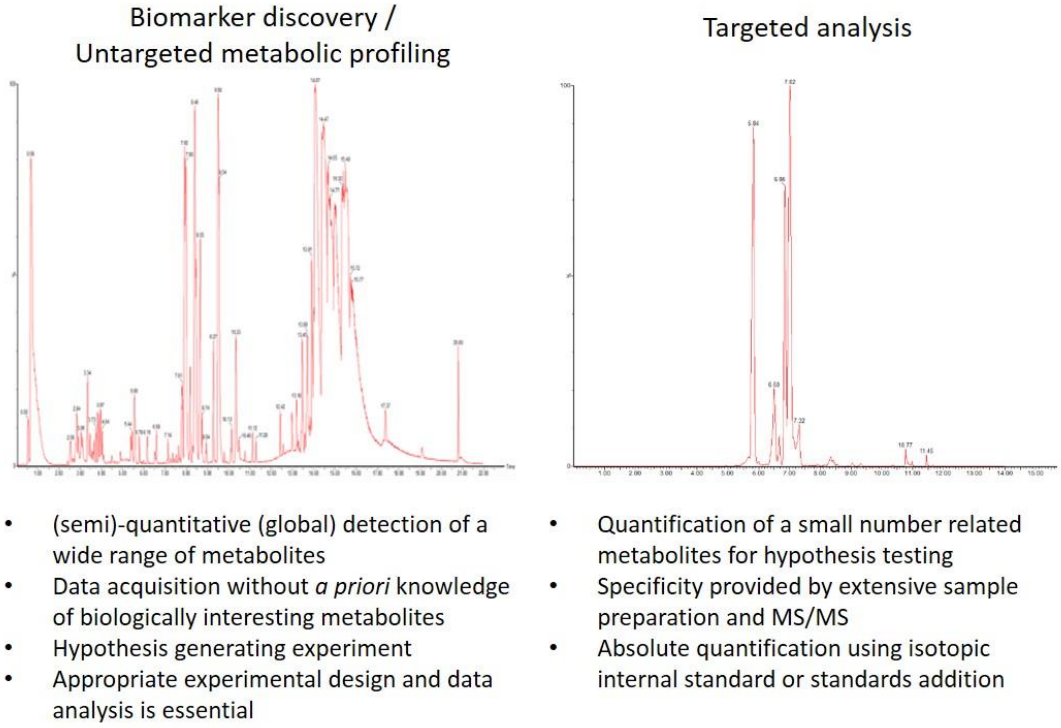


Figure 1.4: Comparison between biomarker discovery (or untargeted metabolic profiling) and targeted analysis in metabolomics. Modified from Dunn *et al*. 2011 (Dunn et al. 2011b).

The use of both of untargeted biomarker discovery and targeted metabolomics methods are useful in clinical metabolomics, and typically progress following these steps: (1) discovery study, (2) technical study validation, and (3) clinical validation. The first two steps are usually performed on matched case-control samples, with a samples number from 15-100s for each group from two independent populations. However, the third step is usually performed on a cross-section of the complete population to test the potential biomarkers using larger numbers of samples,

typically in the 1000s. Such experiments can be difficult to implement without the necessary financial resources, or access to large number of samples.

1.4.4.1. Biomarker discovery or untargeted metabolic profiling

The first approach employed in metabolomics studies is usually untargeted metabolic profiling, which allows for the generation of hypotheses by comparing the levels of detected metabolites between different sample groups without bias (Patti et al. 2012). In the context of clinical metabolomics, untargeted metabolite profiling is employed to determine what are the specific pathophysiological changes between "healthy" and "disease" sample groups. In such studies, the aim is to obtain valid data on as many classes of metabolites as possible, and on metabolites involved with several metabolic pathways across the metabolic network, without *a priori* information (Dunn et al. 2011b). In addition, untargeted profiling of specific sub-classes of metabolites has been reported in recent years. One of the most studied sub-metabolome is focused on lipids, and is named lipidomics (Watson 2006; H.-C. Lee and Yokomizo 2018; Shevchenko and Simons 2010). Whether the full metabolome or an enriched subset of the metabolites are studied, the data generated by untargeted metabolomics methods are then analysed, and once put into the biological context, allow for the discovery of novel hypotheses into the pathophysiological mechanisms, and can highlight potential biomarkers. Sample preparation is generally kept to a limited number of steps in order to prevent metabolite loss. Untargeted metabolic profiling necessitates the use of highly sensitive analytical platforms capable of rapid high-throughput detection of metabolites from a wide range of chemical classes in biological samples, such as MS. The untargeted experiments presented in this thesis were performed on a Waters ACQUITY ultra-high performance liquid chromatography coupled to a quadrupole time-of-flight mass spectrometer (UPLC-Q-TOF-MS) platform (**Figure 1.5**). The data are obtained as the relative intensity for each detected feature, and not quantification of concentrations. Untargeted metabolomic analyses produce large

and complex data sets, which necessitates the use of specific software as manual data treatment is not feasible (**section 1.4.3.2**).

Untargeted metabolomics has, however, led to the discovery of new insights into disease pathophysiology by identifying altered metabolic pathways, and elevated or lowered levels of specific metabolites (Patti et al. 2012). The enzymes involved in the chemical reactions for the biosynthesis of metabolites, or that use them as a substrate in a biochemical reaction, can be targeted by new drugs, and can lead to novel drug treatments (David S Wishart 2016). In recent years, metabolomics techniques have become crucial tools in drug discovery (David S Wishart 2016), and several metabolomics studies have shown metabolic alterations detected in pathologies, such as Alzheimer's diseases (Cai et al. 2012), frailty (Picca et al. 2019), atherosclerosis (J. M. Brown and Hazen 2015), cardiovascular disease (E. Smith et al. 2019), or autism and schizophrenia (Shaw 2010), which have led to new hypotheses on their aetiologies and new drug targets to be tested.



Figure 1.5: Photo of the analytical platform used for the untargeted experiments presented in this thesis. Photo shows a Waters ACQUITY ultra-high performance liquid chromatography coupled to a Synapt G2-S quadrupole time-of-flight mass spectrometer (UPLC-Q-TOF-MS).

1.4.4.2. Targeted analysis

The second analytical approach is used to test the hypotheses generated by untargeted metabolic profiles experiments, and try to validate them in independent samples. Targeted metabolomics analyses are used to obtain with high specificity, precision and accuracy, the (absolute) quantification of the concentrations of a small number of metabolites – generally less than 20 – related to a specific biological pathway or function, or a specific chemical class (Dunn et al. 2011b). Sample preparation and analytical separation can be more complex in order to further separate the metabolites of interest from all others in the sample matrix. Typically, tandem mass spectrometry (MS/MS) is coupled to chromatographic separation, and accuracy is measured by using appropriate internal standards, such as isotopically labelled analogues of the targets.

The most commonly used platform for such analyses is liquid chromatography coupled to triple quadrupole mass spectrometer (LC-TQ-MS), such as the Waters UPLC-TQ-Ms platform used for the experiment presented in this thesis (**Figure 1.6A**). Such platforms allow for multiple reaction monitoring (MRM), also called selected reaction monitoring (SRM) (H.-C. Lee and Yokomizo 2018). Targeted metabolites are selected by their specific ionised mass to charge ratio (m/z) in the first quadrupole (Q1) and are called precursor ions (**Figure 1.6B**). Following this selection, fragmentation of the precursor ion is induced by collision-induced dissociation, in the second quadrupole (Q2). Finally, ions that produce preselected fragments (or product ions) are re-selected in the third quadrupole (Q3) before detection.

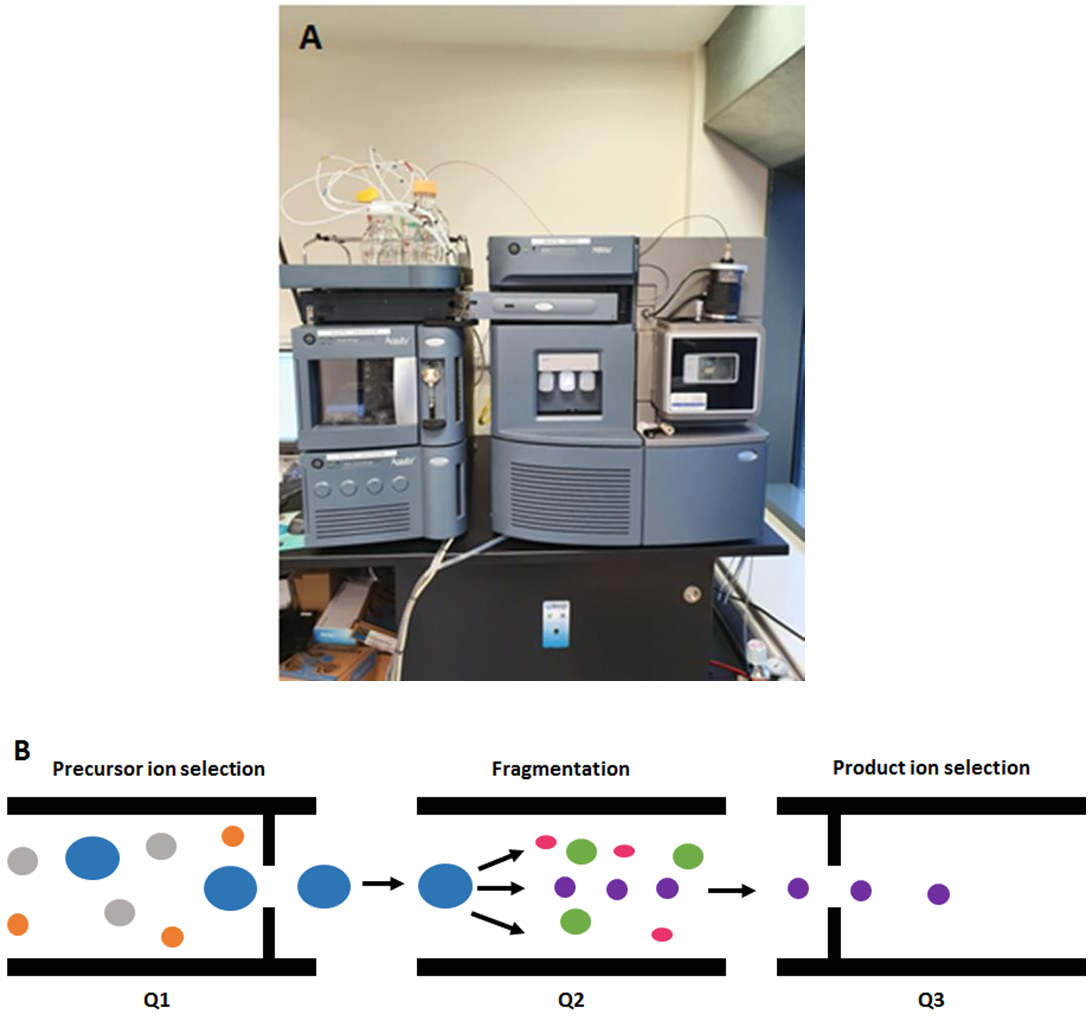


Figure 1.6: Targeted mass spectrometry. Photo A shows a Waters ACQUITY ultra-high performance liquid chromatography coupled to a Xevo triple quadrupole mass spectrometer (UPLC-TQ-MS) used for the targeted experiments describes in this thesis. B shows a schematic of a triple quadrupole mass spectrometer used for targeted analysis and multiple reaction monitoring.

A variation of the targeted approach has recently been developed. This strategy is employed using triple quadrupole mass spectrometers, and requires *a priori* knowledge of which metabolic pathways might be altered in the studied disease (Dunn et al. 2011b). These experiments were developed to obtain a higher specificity, precision and accuracy than untargeted, the (semi-)quantitative concentrations of up to 400 metabolites (Sabatine et al. 2005; Lu et al. 2008) - as opposed to obtaining the concentration of less than 20 in targeted analysis. With this approach, the

metabolites targeted are part of various chemical classes and metabolic pathways, and allow for a wide coverage of the metabolome. For example, such an approach could focus on a specific subset of lipids, such as phosphatidylcholines (PC). Sample preparation would focus on extracting lipids, and the analysis would be performed on a TQ-MS platform, optimised to select and detect only PC lipids. Such experiments are, however, limited, as they require the use of commercially available standards, which means they will only target known metabolites (Dunn et al. 2011b).

1.4.5. Challenges in clinical metabolomics

1.4.5.1. Challenges with sample collection and storage

As metabolites are intermediate molecules of the metabolism, and metabolite flux can be measured in seconds, it is crucial that quenching of the metabolism is performed at time of sampling (e.g. freezing, **section 1.4.3.1**), in order to obtain a sample that accurately reflects the metabolome at time of sample collection (Dunn et al. 2011b; Wood and Cebak 2018). In addition, in clinical metabolomics, and when samples are collected as part of a multicentre study, it is crucial to ensure that collection of samples, their storage and transport follow the same operating procedures standards across all participating centres. The SCreening fOr Pregnancy Endpoints study (SCOPE) birth cohort (Louise C Kenny et al. 2014) is such an example, and samples from this study were used for most of the experiments presented in this thesis.

Furthermore, appropriate storage of the samples prior to their analysis is also crucial to obtain good data quality, especially if samples are expected to be stored for several years. In the case of blood and urine, samples are stored at ultra-low temperatures of -80 °C until time of sample preparation (Villas-Boas et al. 2007). It is also important to minimise the number of freeze/thaw cycles of the samples, as this can lead to changes in thermodynamic properties of the samples, such as pH, viscosity or solubility (Mullen and Critser 2007). The INFANT Centre where all the experiments presented in this thesis were performed, is a biobank facility that follows and exceeds the requirements of the International Society for Biological and

Environmental Repositories (ISBER) best practices. In addition, the biobank adheres to the Molecular Medicine Ireland Guidelines for Standardised Biobanking and OECD Guidelines for Human Biobanks and Genetic Research Databases. The protocols developed by INFANT have contributed to international biobanking standards development (ISO TC 276).

1.4.5.2. Challenges with sample size and power

In clinical metabolomics, and especially when using samples taken from human participants, a major difficulty faced when performing statistical analysis of metabolomics data is that the heterogeneity and differences in the metabolomes of "healthy" and "diseased" samples are often subtle, and can be altering only a small percentage of the metabolome (Dunn et al. 2011b). These subtle differences are all the more difficult to highlight if the sample size of the groups compared are not enough to appropriately power the study and detect statistically significant differences between the groups (Sawilowsky 2009). The SCOPE study was established for this reason, and to be able to collect thousands of samples. However, pregnancy adverse outcomes have low prevalence, and especially in the low-risk pregnant population, which was recruited into SCOPE cohort. Challenges with sample size are also addressed in the discussion (**section 7.4**).

To help address this challenge, various statistical algorithms can help address studies with small sample size, such as machine learning methods (Random Forests, support vector machine) (T. Chen et al. 2013; Mahadevan et al. 2008), or Bayesian methods (Smyth 2004). The latter has been used throughout this thesis. Bayesian methods interpret each parameter as a random distribution and use a ranking of the metabolites according to their ability to predict the outcome.

Regardless of the statistical methods chosen, the complexity of metabolomics datasets leads to high rates of false discovery, especially when using supervised methods and powerful algorithms that can easily find random associations (Dunn et al. 2011b). This is especially true with untargeted MS that can detect several thousands of metabolites per sample. Statistical analysis must include multiple

testing correction methods, such as the Benjamini-Hochberg correction where the p-values are adjusted based on the number of calculations performed, thus reducing false discovery rates (Benjamini and Hochberg 1995). For this reason, an emphasis is also placed on the importance of validation.

1.4.5.3. Challenges with biomarker validation

Biomarker validation studies present several limitations, mainly that data are not easily compared across different platforms due to inter-laboratory instrument variations (Griffiths et al. 2010). However, it is possible to minimise technical variations by using internal (added to the samples before extraction) or external (added to the samples after extraction) standards. Another problem is metabolite identification. Biological inference from metabolic features detected necessitate their chemical identification. To date, metabolite identification is still one of the main challenges of the metabolomics field, as thousands of metabolites are detected in each experiment (Dunn et al. 2013). Indeed, absolute identification is achieved by comparing empirical data to standards analysed in the same conditions. This can be onerous and extremely time consuming, especially when hundreds of metabolites of interest have been detected, many of which are unknown molecules requiring *de novo* structural characterisation. For metabolomics studies, this identification process is still manual or semi-automated at best. In the case of studies using electrospray ionisation (ESI), such as those described in this thesis, a wide range of ions can be detected, and include fragments, isotope, multiple charged peaks, and adducts (M. Brown et al. 2009; Draper et al. 2009). Nevertheless, *in silico* identification through library and database searches can help identification to some extent.

In addition to technical limitations, there are clinical challenges for biomarker validation, especially access to samples from an independent cohort. Indeed, as the overarching goal of biomarker validation is to be applicable to the worldwide population, being able to test and validate the biomarker candidates in samples from a different set of sample than the ones used for discovery study, and with higher

numbers of samples per group is the ideal (Dunn et al. 2011b). However, this can prove difficult to do as some clinical studies are focused on very specific diseases, and samples from a reliable source can be quite rare. In addition, the cost of analyses performed in cohorts of patients can prove to be onerous. As detailed in chapter 5, and to address this challenge, we are one of the only research group that has undertaken biomarker validation of spontaneous preterm birth in longitudinal samples.

Despite all these challenges, metabolomics biomarker validation can lead to successful screening tests. An example is the PrePsia^{TM} developed by Metabolomics Diagnostics (www.metabolomicdiagnostics.com) to detect women at risk of developing preterm pre-eclampsia. This test is performed on blood samples collected at 15 weeks of gestation and using a combination of metabolites and clinical data. This test is in the late stages of research, soon to be commercialised.

1.5. Hypotheses and aims

Overall Objective and rationale: To use metabolomics techniques to study the pathophysiology of adverse pregnancy outcomes**.** Pregnancy complications - such as small for gestational age babies, preterm birth, or pre-eclampsia - are complex, caused by multiple factors, and are critical perinatal health issues. However, to date their pathophysiology remains unclear, and no accurate predictive test has been developed. Clinical metabolomics techniques allow for the sensitive measure of metabolites, which are a direct reflection of the phenotype. Therefore, I used metabolomics techniques to help decipher the molecular pathways implicated in maternal plasma and urine samples of women who later delivered a small for gestational age (SGA) newborn, or spontaneous preterm birth (sPTB). In addition, I applied metabolomics techniques to analyse the effect of a potential drug treatment on the plasma metabolome of a rat model of pre-eclampsia.

Hypothesis 1: The metabolome of women who developed pregnancy adverse outcomes are significantly altered compared to that of women who had an uncomplicated pregnancy.

Aim 1: To investigate the predictive accuracy of metabolomics as a tool for prediction of SGA by undertaking a systematic review of the literature. ***Chapter 2***

Aim 2: To investigate the plasma and urine metabolome in SGA at 20 weeks of gestation in comparison to controls. ***Chapter 3***

Aim 3: To identify the altered lipids associated with sPTB at 20 weeks of gestation in plasma samples. ***Chapter 4***

Aim 4: To identify potential lipid biomarkers of sPTB in plasma samples in Cork and Auckland populations at 15 weeks of gestation, and to validate biomarker candidates them at 15, 20 weeks of gestation and at time of delivery. ***Chapter 5***

Hypothesis 2: The metabolome of a rat model of pre-eclampsia is altered after the treatment with a potential therapeutic agent for pre-eclampsia.

Aim 1: To determine the effects of L-Ergothioneine, a potential drug treatment for pre-eclampsia, on the plasma metabolome of a rat model of pre-eclampsia. ***Chapter 6***

Chapter 2:

Examining the predictive accuracy of metabolomics for small for gestational age babies: a systematic review.

Leite DFB, Morillon A-C, *et al*. Examining the predictive accuracy of metabolomics for small-for-gestational-age babies: a systematic review. BMJ Open 2019;9:e031238. doi:10.1136/bmjopen-2019-031238 (Leite et al. 2019)

2.1. Abstract

Introduction:

To date, there is no robust enough test to predict small-for-gestational-age (SGA) infants, who are at increased lifelong risk of morbidity and mortality.

Objective:

To determine the accuracy of metabolomics in predicting SGA babies and elucidate which metabolites are predictive of this condition.

Data sources:

Two independent researchers explored 11 electronic databases and grey literature in February 2018 and November 2018, covering publications from

1998 to 2018. Both researchers performed data extraction and quality assessment independently. A third researcher resolved discrepancies.

Study eligibility criteria:

Cohort or nested case–control studies were included which investigated pregnant women and performed metabolomics analysis to evaluate SGA infants. The primary outcome was birth weight <10^{th} centile—as a surrogate for fetal growth restriction—by population-based or customised charts.

Study appraisal and synthesis methods:

Two independent researchers extracted data on study design, obstetric variables and sampling, metabolomics technique, chemical class of metabolites, and prediction accuracy measures. Authors were contacted to provide additional data when necessary.

Results:

A total of 9,181 references were retrieved. Of these, 273 were duplicate, 8,760 were removed by title or abstract, and 133 were excluded by full-text content. Thus, 15 studies were included. Only two studies used the fifth centile as a cut-off, and most reports sampled

second-trimester pregnant women. Liquid chromatography coupled to mass spectrometry was the most common metabolomics approach. Untargeted studies in the second trimester provided the largest number of predictive metabolites, using maternal blood or hair. Fatty acids, phosphosphingolipids and amino acids were the most prevalent predictive chemical subclasses.

Conclusions and implications:

Significant heterogeneity of participant characteristics and methods employed among studies precluded a meta-analysis. Compounds related to lipid metabolism should be validated up to the second trimester in different settings.

PROSPERO registration number CRD42018089985.

2.2. Introduction

Fetal growth restriction (FGR) and small for gestational age (SGA) infants are major concerns in modern obstetrics (Galan and Grobman 2015; Figueras and Gratacos 2014; Gordijn et al. 2016). SGA is commonly used as a proxy for FGR (Bukowski et al. 2014), despite the subtle differences between these two pathological conditions. The prevalence of both varies according to criteria applied and on the population and setting, although it reaches as much as 25 % in low and middle-income countries (A. C. Lee et al. 2013). SGA newborns may have adverse health effects, such as stillbirth (Bukowski et al. 2014), perinatal asphyxia (Mendez-Figueroa et al. 2016), impaired neurodevelopment (D. Sharma et al. 2016a), and increased cardiovascular risk (Barker et al. 1993; Balasuriya et al. 2018). To date, there are no robust prediction tools for SGA using clinical factors (Goto 2016; Pay et al. 2016), ultrasound data (Parry et al. 2017; Roma et al. 2015), or placental biomarkers (Conde-Agudelo A et al. 2013).

For hypothesis generating or validation purposes, metabolomics is a novel area of biomarker, discovery, development and clinical diagnostics in translational medicine (Xia et al. 2013; Patti et al. 2012). Metabolomics is the study of all metabolites (Xia et al. 2013; Patti et al. 2012; Horgan et al. 2011) in a given sample, i.e. low molecular weight compounds (50-2000 Da) that are intermediates of biochemical reactions and

metabolic pathways, considered to directly reflect cellular activity and phenotype (Xia et al. 2013; Patti et al. 2012). Recent studies have evaluated the pathophysiology (Sulek et al. 2014; Favretto et al. 2012; Sanz-Cortés et al. 2013; Monier et al. 2016) of SGA with metabolomics. However, little is known about the potential of metabolomics to identify predictive compounds of SGA.

Since metabolomics can identify multiple metabolites from low volume samples, and create a model from a collection of these samples (Xia et al. 2013), it is a promising technology for hypothesis generation in a heterogeneous condition such as SGA. The prediction of SGA in pregnancy would help refer women to specialized care facilities, improving maternal and neonatal outcomes (Verlijsdonk et al. 2012; Leite et al. 2018).

In this context, our review question was 'What is the accuracy of metabolomics for predicting FGR?'. The main objective of this systematic review was to assess the accuracy of metabolomics techniques in predicting SGA. As a secondary aim, we intended to determine which metabolites are predictive of this condition.

2.3. Methods

The protocol for this systematic review was published previously (Leite et al. 2018). This study follows international guidelines for transparency (International Prospective Register of Systematic Reviews) and respects the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (Moher et al. 2009).

2.3.1. Literature search strategy

Two independent researchers (DFBL and A-CM) assessed 11 electronic databases (PubMed, EMBASE, Latin American and Caribbean Health Sciences Literature, Scientific Electronic Library Online, Health Technology Assessment, Database of Abstracts of Reviews of Effects, Aggressive Research Intelligence Facility, Cumulative Index of Nursing and Allied Health Literature, Maternity and Infant Care, Scopus, and Web of Science) and grey literature. There were no limits or language constraints;

the search strategy covered published documents between 1998 and 2018. The keywords ‘small for gestational age’, ‘metabolomics’, ‘prediction’ and ‘antenatal’, and variations of each, were combined with Boolean operators depending on each database requirements. The full EMBASE literature search was as follows: (‘fetal growth retardation’ OR ‘fetal growth restriction’ OR ‘intrauterine growth restriction’ OR ‘intrauterine growth retardation’ OR ‘small for gestational age’) AND (‘metabolomic*’ OR ‘metabonomic*’ OR ‘metabolit* ‘H NMR’ OR ‘proton NMR’ OR ‘proton nuclear magnetic resonance’ OR ‘liquid chromatogra*’ OR ‘gas chromatogra*’ OR ‘UPLC’ OR ‘ultra-performance’ OR ‘ultra performance liquid chromatograph*’) AND (‘pregnan*’ OR ‘antenat*’ OR ‘ante nat*’ OR ‘prenat*’ OR ‘pre nat*’) AND (‘screen*’ OR ‘predict*’ OR ‘metabolic profil*’). Please check online **Table 2.1** for more details.

Table 2.1: Detailed literature search strategy.

1 fetal growth retardation
2 fetal growth restriction
3 intrauterine growth restriction
4 intrauterine growth retardation
5 small for gestational age
6 #1 OR #2 OR #3 OR #4 OR #5
7 metabolomic*
8 metabonomic*
9 metabolit*
10 H NMR
11 proton NMR
12 proton nuclear magnetic resonance
13 liquid chromatogra*
14 gas chromatogra*
15 UPLC
16 ultra-performance liquid chromatograph*
17 ultra performance liquid chromatograph*
18 #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17
19 pregnan*
20 antenat*
21 ante nat*
22 prenat*
23 pre nat*
24 #19 OR #20 OR #21 or #22 OR #23
25 screen*
26 predict*
27 metabolic profil*
28 #25 OR #26 OR #27
29 #6 AND #18 AND #24 AND 28

2.3.2. Outcomes and subgroup analysis

The primary outcome was SGA, as a surrogate for FGR and defined as birth weight <10th centile, by population-based or customised charts. The secondary outcomes were birth weight ≤5th or ≤3rd centile.

The intended subgroup analysis comprised the type of metabolomics technique applied (nuclear magnetic resonance, NMR; gas or liquid chromatography coupled with mass spectrometry, GC-MS or LC-MS, respectively); maternal health status before pregnancy (women with vs without any chronic health condition); type of SGA suspected during pregnancy (early vs late SGA); and type of pregnancy (singleton vs multiple pregnancy).

2.3.3. Selection criteria of studies, data collection and analysis

Cohort or case–control studies were included if maternal samples were collected before the clinical diagnosis of SGA, if any metabolomics technique was applied and if the results of SGA were presented. Articles presenting data from the same research project but analysing distinct metabolites or showing data from different countries were included. Studies were excluded (1) according to study design; (2) if they had not applied any metabolomics technique; (3) if they were only experimental studies; (4) if it was not possible to extract data on SGA; or (5) if they presented duplicate data, in which case the most complete publication was included for final analysis.

Two researchers (DFBL and A-CM) independently selected studies, extracted data and discussed discrepancies. One additional reviewer (EFMJ or RTS) helped to decide, by majority, when no consensus was reached.

Piloted standardised forms were applied for data extraction, including pregnancy characteristics and experimental details. The Human Metabolome Database (HMDB) (David S. Wishart et al. 2018) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) were used for matching chemical class and metabolic pathways of each metabolite, respectively.

2.3.4. Risk of bias and applicability concerns assessment concerns regarding applicability

Two researchers (DFBL and A-CM) independently evaluated individual studies using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (Whiting et al. 2011). One of the third reviewers (EFMJ or RTS) helped in decision-making when no consensus was achieved.

Each study was classified as high, low or unclear risk of bias in four domains (patient selection, index test, reference standard, and flow and timing), and as high, low or unclear concerns regarding applicability in the first three domains. We did not consider two signalling questions ('Was a case-control design avoided?' and 'Was there an appropriate interval between the index test and reference standard?'). The nested case–control design was an inclusion criterion, and maternal samples should have been collected during pregnancy, that is, before the SGA diagnosis. Studies were considered 'low risk', for example, (1) if pregnancy or neonatal complications were not excluded in just one group of participants or data on participant selection had been provided; (2) if methods for sample preparation and interpretation were standardised or metabolite threshold was defined before the experiments (for targeted analysis); (3) if the adequacy and reasons for choosing the reference birthweight chart had been explained; or (4) if large-for-gestational-age babies had been excluded from the final comparative analysis.

2.3.5. Data synthesis

A quantitative summary of data was performed when any predictive accuracy measures could be extracted. Authors were contacted to provide additional information, when necessary. However, only Delplancke et al (Delplancke et al. 2018) replied. The estimation of likelihood ratios and hierarchical summary receiver operator characteristic curve (Rutter and Gatsonis 2001) was planned, as well as assessment of heterogeneity and publication bias (Deeks et al. 2005). However, due to lack of data, a meta-analysis could not be performed.

2.3.6. Patient and public involvement

There was no patient or public involvement in conducting this systematic review.

2.4. Results

2.4.1. Literature search characteristics

The literature search for this systematic review was performed in February 2018 and rerun in November 2018. A total of 9,181 references were retrieved (**Figure 2.1**). After the removal of duplicate records (n=273), title and abstract screening, and analysis of the remaining 148 full-text articles, 15 articles were included (Horgan et al. 2011; Sulek et al. 2014; Delplancke et al. 2018; Costet et al. 2011; Ertl et al. 2012; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013; Gong et al. 2018; Morillon et al. 2018). See **Appendix B** for the excluded studies.

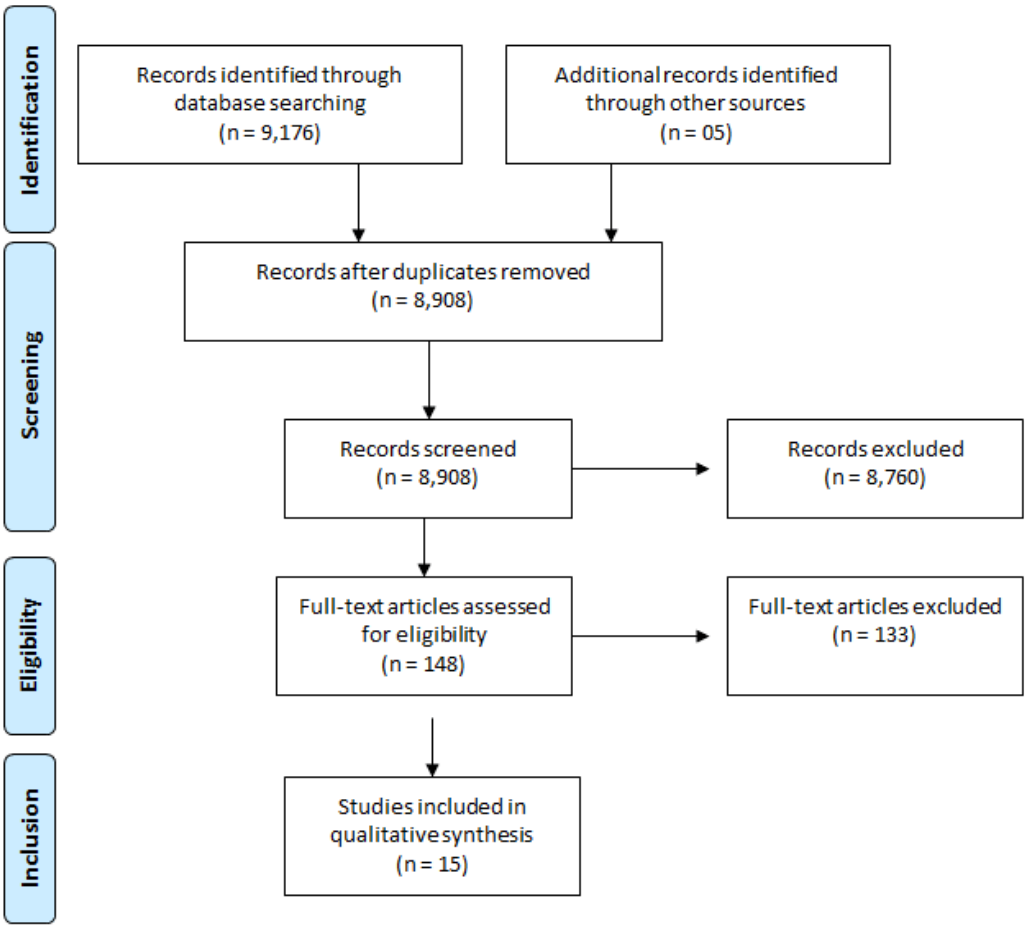


Figure 2.1: PRISMA flow chart of study identification, screening and selection. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses. From Moher D, et al (Moher et al. 2009) For more information, visit www.prisma-statement.org.

2.4.2. Characteristics of the included studies

The characteristics of the included studies are shown in **Table 2.2**. The prevalence of GA ranged from 7.3 % (Grandone et al. 2006) to 21.5 % in cohort studies (Delplancke et al. 2018). There were no studies using birth weight ≤3rd centile to define SGA. The time interval between initial participant enrolment and publication varied from 3 (Horgan et al. 2011) to 54 years (Gernand et al. 2013), although these data were unclear in 38 % of the reports (Sulek et al. 2014; Delplancke et al. 2018; Ertl et al. 2012; Grandone et al. 2006; Ong et al. 2016). In nested case– control studies, participants were matched by maternal age (Horgan et al. 2011; Sulek et al. 2014;

Luthra et al. 2018; Morillon et al. 2018), ethnicity (Horgan et al. 2011; Sulek et al. 2014; Morillon et al. 2018), parity (Luthra et al. 2018), body mass index (Horgan et al. 2011; Sulek et al. 2014; Morillon et al. 2018) or infant gender (Luthra et al. 2018; Sulek et al. 2014).

Participant characteristics varied between studies. Regarding gestational age at assessment, samples were collected in the second trimester in half of the studies (Horgan et al. 2011; Sulek et al. 2014; Grandone et al. 2006; Kiely et al. 2016; Ong et al. 2016; van Eijsden et al. 2008a; Morillon et al. 2018). In three reports, women were assessed at least twice (Choi et al. 2016; Luthra et al. 2018; Gong et al. 2018). In one study, maternal blood was drawn either in the first or second trimester (Gernand et al. 2013), and in another three studies only samples from the third trimester were considered (Delplancke et al. 2018; Y. Wang et al. 2016; Gong et al. 2018). In the latter case, maternal hair was divided according to length, allowing evaluation of second-trimester and third-trimester metabolites (Delplancke et al. 2018). Studies considering the fifth centile as the cut-off sampled women in the first trimester (Costet et al. 2011; Ertl et al. 2012). Twin pregnancy was a clear exclusion criterion in most studies (Horgan et al. 2011; Sulek et al. 2014; Costet et al. 2011; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Ong et al. 2016; Gernand et al. 2013; Gong et al. 2018; Morillon et al. 2018). Pregnancy aided by assisted reproduction (Sulek et al. 2014; Ong et al. 2016) or women with pre-existing conditions (Horgan et al. 2011; Sulek et al. 2014; Kiely et al. 2016; Ong et al. 2016; Morillon et al. 2018) were also excluded, although these data were incompletely reported (Delplancke et al. 2018; Ertl et al. 2012; Y. Wang et al. 2016; Luthra et al. 2018; Ong et al. 2016; Gong et al. 2018). When both nulliparous and multiparous women were enrolled, there was no data analysis according to parity. Half of the studies considered term deliveries exclusively (Sulek et al. 2014; Delplancke et al. 2018; Y. Wang et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013; Gong et al. 2018), and the remaining studies did not differentiate results according to gestational age at birth.

Regarding clinical risk factors for SGA, only one paper mentioned a history of SGA, but findings were not adjusted for this variable (Ertl et al. 2012). All studies, except one (Delplancke et al. 2018), cited participants' smoking status. The rate of smoking

habit ranged from 2.4 % (Sulek et al. 2014) to 47.5 % (Gernand et al. 2013). It is important to note that Gernand *et al* (Gernand et al. 2013) analysed samples from women recruited between 1959 and 1965, when smoking while pregnant was encouraged, which explains the high rate of smoking participants. The duration of smoking or any differences in birth weight (absolute measures or centiles) were not clearly stated. Although more prevalent in SGA pregnancies, the results did not change with this variable control (Costet et al. 2011; Ertl et al. 2012; Kiely et al. 2016; Ong et al. 2016; Gernand et al. 2013). Only Gong *et al* (Gong et al. 2018) mentioned the suspicion of SGA in pregnancy, exhibiting decreasing abdominal circumference growth velocity between 20 and 36 weeks. However, on final analysis, these babies were grouped with infants not suspected during pregnancy.

Table 2.2: Main characteristics of included studies.

Authors, year	Country, year of participants' enrolment	Study design	Affected/ non-affected	Gestational age at assessment	Type of pregnancy	Parity	Birthweight curve
Outcome: SGA <5th centile							
Costet N *et al*, 2011	France, 2002-2006 (PELAGIE Cohort)	Nested case-control	134/ 399	11w	Single pregnancy	Nulliparous and parous women, unclear proportions	Customized curve
Ertl R *et al*, 2012	United Kingdom [a]	Nested case-control	150/ 1,000	11^{+0}-13^{+6}w	Unclear	55,3% nulliparous in SGA group, 48.1% nulliparous in control group	Population-based charts
Outcome: SGA <10th centile							
Grandone E *et al*, 2006	Italy [a]	Cohort	31/ 393	17.1 ± 1.2w[b] (mean)	Single pregnancy; no maternal pre-existing conditions	Unclear	Population-based charts
van Eijsden M *et al*, 2008	Netherlands, 2003-2004 (ABCD Study)	Cohort	429/ 3275	13.5 ± 3.3w (mean)	Term deliveries, no diabetes or hypertension	57.6% nulliparous	Population-based charts
Horgan *RP et al*, 2011	Australia, 2008-2011 (SCOPE Cohort)	Nested case-control	40/ 40	14-16w	Single pregnancy; no other pregnancy complications	Nulliparous	Customized curve

Gernand AD *et al*, 2013	United States, 1959-1965 (Collaborative Perinatal Project)	Nested case-control	395/ 1751	≤26w	Single pregnancy; term deliveries	Parous women	Population-based charts
Sulek K *et al*, 2014	Singapore [a] (GUSTO Study)	Nested case-control	41/ 42	26-28w	Single pregnancy; term deliveries; no maternal pre-existing conditions	Nulliparous and parous women, unclear proportions	Population-based charts
Choi R *et al*, 2016	South Korea, 2012-2013	Cohort	39/ 217	1st, 2nd or 3rd trimester	Single pregnancies	Nulliparous and parous women, unclear proportions	Population-based charts
Kiely ME *et al*, 2016	Ireland, 2008-2011 (SCOPE Cohort)	Cohort	190/ 1578	14-16w	Single pregnancy; no maternal pre-existing conditions	Nulliparous	Customized curve
Ong YL *et al*, 2016	Singapore [a] (GUSTO Study)	Cohort	83/ 827	26-28w	Single pregnancy; no maternal chronic illness	43.5% nulliparous	Population-based charts
Wang Y *et al*, 2016	Taiwan, 2000-2001 (Taiwan Maternal and Infant Cohort Study)	Cohort	35/ 188	3rd trimester	Unclear; term deliveries	48% nulliparous	Population-based charts
Delplancke TDJ *et al*, 2018	New Zealand [a]	Cohort	20/ 73	34-37w	Unclear; term deliveries	Unclear	Customized curve

Luthra G *et al*, 2018	United States, 2010-2012 (TIDES Study)	Nested case-control	53/ 106	1st and 2nd trimester	Single pregnancies; term deliveries	60% nulliparous	Customized curve
Gong S *et al*, 2018	United Kingdom, 2008-2012 (POP study)	Nested case-control	162/259	36w	Single pregnancies; term deliveries	Nulliparous	Customized curve
Morillon A-C *et al*, 2018	2008-2011 (SCOPE Study)	Nested case-control	40/40	20w	Single pregnancies	Nulliparous	Customized curve

[a] Unclear period of participant recruitment. [b] Mean for all study participants.

2.4.3. Subgroup analysis

Due to unavailable data, the only subgroup analysis performed was related to the metabolomics approach applied (**Table 2.3**). There was no mention of adherence to metabolomics reporting data guidelines. LC-MS was the leading technique used. Three studies have investigated metabolites related to environmental exposure, from contaminated water(Costet et al. 2011), consumer products (Y. Wang et al. 2016), or pesticides (Morillon et al. 2018), while others have analyzed endogenous compounds (Ertl et al. 2012; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Ong et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013). Only Luthra *et al* conducted a biomarker validation study (Luthra et al. 2018), while Gong *et al* (Gong et al. 2018) chose to analyze the top ten statistically different metabolites according to infant sex.

Maternal blood was the most common biological sample analyzed by LC-MS in all studies (Horgan et al. 2011; Ertl et al. 2012; Choi et al. 2016; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; van Eijsden et al. 2008a; Gernand et al. 2013; Gong et al. 2018), except for one, which used GC-MS (van Eijsden et al. 2008a). Maternal urine was analyzed by NMR (Luthra et al. 2018), GC-MS (Y. Wang et al. 2016) or LC-MS (Morillon et al. 2018). There was only one report using amniotic fluid (Grandone et al. 2006) and two using maternal hair (Sulek et al. 2014; Delplancke et al. 2018), all applying GC-MS. The period of laboratory analysis was rarely specified, which made it impossible to estimate total time of sample storage.

Untargeted studies reported diverse metabolic features. Authors matched the peaks with an in-house library (Sulek et al. 2014; Delplancke et al. 2018), or HMDB-related database (Horgan et al. 2011; Morillon et al. 2018). Horgan *et al* (Horgan et al. 2011) found 785 compounds both in maternal and newborn samples; their predictive model included 19 metabolites (only five could be putatively identified, **Table 2.2**) and used 2nd trimester maternal blood. Sulek *et al* (Sulek et al. 2014) and Delplancke *et al* (Delplancke et al. 2018) prepared and analyzed samples with GC-MS using similar protocols. Sulek *et al* (Sulek et al. 2014) identified 32 statistically different chromatographic features from which they built a predictive model using five metabolites, including two fatty acids (2-methyloctadecanoate and margarate). In

contrast, Delplancke *et al* (Delplancke et al. 2018), identified 198 metabolites, including three fatty acids (margaric, pentadecanoic, and myristic acid) showing significantly higher levels in SGA cases, when 2nd trimester maternal hair segments were studied.

Table 2.3: Subgroup analysis of included studies according to which metabolomics technique was applied.

Authors/ year	Metabolomics Technique	Maternal sample/ Storage temperature	Prediction model*	Targeted compounds	Coefficient of variation/ Limits of quantitation	Predictive compounds	Sensitivity/ Specificity	AUC
Nuclear magnetic resonance								
Luthra *G et al*, 2018	^{1}H-NMR 1D NOESY with pre-saturation and homonuclear 2D *J*-resolved at 300 K Bruker 600 MHz Advance III HD spectrometer	Urine/ -80^{o}C	Targeted	Tyrosine, acetate, formate, trimethylamine	NA	None		
Gas chromatography coupled to mass spectrometry								
Costet *N et al*, 2011	GC-MS Simple head space SPME-Capillary GC	Urine/ -20^{o}C	Targeted	Trichloroacetic acid	<5%/ 0.01mg/L	None	0.1/ 0.93	
Sulek K *et al*, 2014	GC-MS Thermo Trace GC Ultra system coupled to ISQ mass selective detector Capillary GC column: Phenomenex ZB-1701 (30 m x 250 µm id x 0.15 µm with 5 m guard column)	Hair/ -20^{o}C	Untargeted	NA	NA	↓ Lactate ↓ Levulinate ↑2-methyloctadecanate ↑Tyrosine ↓ Margarate		0.998

Delplancke TDJ *et al*, 2018	GC-MS: Agilent 7890B gas chromatograph, capillary column ZB-1701 (30m x 250µm id x 0.15µm with 5m guard column) 5977 A mass spectrometer, electron impact ionisation	Hair/ -20°C	Untargeted	NA	NA	↑ Margaric acid ↑ Pentadecanoic acid ↑ Myristic acid[c]	0.72 0.73 0.73
Liquid chromatography coupled to mass spectrometry							
Grandone E *et al*, 2006	LC-MS/MS triple quadrupole Applera API 3000, TurboIonSpray ionisation	Amniotic fluid/ -80°C	Targeted	Homocysteine	Unclear	↑Homocysteine (1,29µM; 1,05-1,51µM)	
Horgan RP *et al*, 2011	UPLC- MS/MS Thermo Fisher LTQ Orbitrap, ESI	Plasma/ -80°C	Untargeted	NA	NA	Hexacosanedioic acid, diglyceride, lyso-phosphocholine, sphinganine 1-phosphate; sphingosine 1-phosphate[d]	0.90
Ertl R *et al*, 2012	HPLC- MS/MS Shimadzu Prominence HPLC system with a column Phenomenex Luna C8 3 x 50 mm; AbSciex API-5000 triple quadrupole, ESI	Serum/ -80°C	Targeted	25(OH)D_2; 25(OH)D_3	6.3%[a], 6.6%[b] (D_2); 6.5%[a], 7.3%[b] (D_3)/ unclear	↓25,OH,Vitamin D (12.16ng/mL; 8.09-20.54ng/mL)	0.72/ 0.45

Gernand AD *et al*, 2013	LC-MS/MS	Serum/ -20°C	Targeted	25(OH)D_2; 25(OH)D_3	8.2%[a] (D_2) 5.9%[a] (D_3)/ <1ng/mL	None	0.39/ 0.66
Choi R *et al*, 2016	HPLC- MS/MS Waters HPLC system, Applied Biosystems API-4000 MS/MS mass spectrometer	Serum/ -20°C	Targeted	Methylmalonic acid; homocysteine	<10%[a]; <10%[b]/ Unclear	None	
Kiely ME *et al*, 2016	UPLC- MS/MS Waters Acquity UPLS system, Waters Triple Quadrupole TQD mass spectrometer	Serum/ -80°C	Targeted	25(OH)D_2; 25(OH)D_3; 3-epi-25(OH)D_3.	<6%[a]; <5%[b]/ 0.57ng/mL (D_2); 0.26ng/mL (D_3), 0,41ng/mL (epi-D_3)	None	
Ong YL *et al*, 2016	LC-MS/MS Applied Biosystems ThermoHypersil BDS C8 reverse-phase column	Plasma/ Unclear	Targeted	25(OH)D_2; 25(OH)D_3	≤10,3%[a,b]/ <1,6ng/mL	None	0.12/ 0.87
Wang Y *et al*, 2016	LC-MS Agilent HPLC system, Applied Biosystems Sciex API-4000 triple quadrupole mass spectrometer	Serum/ Unclear	Targeted	PFOA; long-chain PFCA	0,83-7,94%[a]; 1,57-24,7%[b]/ 0,07-0,45ng/mL[e]	PFDeA (OR 3,14; 95%CI 1,07-9,19), PFUnDA (OR 1,83; 95%CI 1,01-3,32)[f]	
Gong S *et al*, 2018	LC-MS/MS Shimadzu UK Limited UPLC system, ACE Excel 2 C18-PFP LC-column; Thermo Fisher	Serum/ Unclear	Untargeted	NA		↑N1,N12-diacetylspermine[f]	

	Scientific Exactive orbitrap mass spectrometer					
Morillon A-C *et al*, 2018	UPLC- MS/MS Waters Acquity UPLS system, Waters Synapt G2-S mass spectrometer	Urine/-80oC	Untargeted	NA		None
Others						
van Eijsden M *et al*, 2008	GC-FID Solid phase extraction SPE, Capillary GC	Plasma/ -80°C	Semi-targeted, Lipid extraction	Elaidic, linoleic, alfa-linolenic, eicosatetraenoic, EPA, DPA, DHA DGLA, AA, Adrenic, and Osbond acids	≤2 - 22%[b]/ Unclear	↓ Eicosatetraenoic acid (OR 1,5; 95%CI 1,07-2,11), ↓DPA (OR 1,49; 95% CI 1,06-2,1)

[a]Intra-assay and [b]inter-assay coefficients of variation. [c]These metabolites were found in 2nd trimester hair segments. [d]And more 14 metabolites that could not be identified certain based on chromatographic peak and mass: Phenylacetylglutamine or formyl-N-acetyl-5-methroxykynurenamine; leucyl-leucyl-norleucine or sphingosine 1-phosphate; cervonyl carnitine and/or 1-alpha,25-dihydroxy-18-oxocholecalciferol; (15Z)-tetracosenoic acid or 10,13-dimethyl-11-docosyne-10,13-diol or trans-selacholeic acid; pencosenoic acid or cyclohexyl acetate or octanoic acid or methyl-heptenoic acid or 4-hydroxy-2-octenal or DL-2-aminooctanoic acid or 3-amino-octanoic acid; hydroxybutyrate or hydroxy-methylpropanoate or methyl methoxyacetate; lysophosphocoline and phosphocoline (more than 10 hits); phosphocoline (more than 20 hits); phosphocoline or ubiquinone-8; acetylleucil-leucil-norleucinal or oleoylglycerone phosphate or LPA(0:0/18:2(9Z,12Z)) or 1-16:1lysoPE or phosphocoline(O-11:1(10E)/2:0) or (3s)-3,4-Di-N-hexanoyloxybutyl-1-phosphocoline or N-(3-hydroxy-propyl) arachidonoyl amine or N-methyl N-(2-hydroxy-ethyl) arachidonoyl amine or similar; lysophosphocholine (16:1) or cervonyl carnitine; preganediol-3-glucuronide or 3-alpha,20-alpha-dihydroxy-5-beta-pregnane-3-glucuronide; 6-hydroxyshingosine or (4OH,8Z,t18:1) sphingosine or 15-methyl-15-prostaglandin D2 or 15-R-prostaglandin E2 methylester. [e]Values for all studied metabolites. [f]Predictive compounds only for female babies.

AUC: area under the receiver operating characteristic curve; ^{1}H-NMR: hydrogen nuclear magnetic resonance; NOESY: nuclear Overhauser effect spectroscopy; GC-MS: gas chromatography coupled to mass spectrometry; SPME: solid phase micro extraction; LC-MS: liquid chromatography coupled to mass spectrometry; UPLC: ultra-performance liquid chromatography; ESI: Electrospray ionisation; FID: flame ionisation detection; PFOA: perfluorooctanoic acid; PFCA: perfluorocarboxylic acid; PFDeA: perfluorodecanoic acid; PFUnDA: perfluoroundecanoic acid; EPA: eicoisapentaenoic acid; DPA: docosapentaenoic acid; DHA: docosahexaenoic acid; DGLA: dihomo-gama-linolenic acid; AA: arachidonic acid; OR: odds ratio; CI: confidence interval; NA: not applicable.

2.4.4. Analysis of identified metabolites

The identified compounds refer to eleven HMDB chemical classes. Fatty acids (Sulek et al. 2014; Delplancke et al. 2018; van Eijsden et al. 2008a) comprised the most prevalent chemical class, followed by amino acids (Sulek et al. 2014; Grandone et al. 2006) and phosphosphingolipids (Horgan et al. 2011) (**Table 2.4**).

A total of 5,974 women were assessed for vitamin D status. Results were presented as total vitamin D (Ertl et al. 2012; Kiely et al. 2016; Ong et al. 2016; Gernand et al. 2013), although vitamin D2, D3 or 3-epi-25(OH)D3 (Kiely et al. 2016) metabolites were measured. Results were stratified according to season of maternal sampling or latitude. Either <15ng/mL (<37.5nmol/L) (Gernand et al. 2013) or <20ng/mL (<50nmol/L) (Ertl et al. 2012; Kiely et al. 2016; Ong et al. 2016) levels characterized vitamin D deficiency, but were statistically different in SGA pregnancies only in the 1st trimester (Ertl et al. 2012). Horgan *et al* found a metabolite that could represent a vitamin D derivative, but it was only predictive in combination with 18 other compounds; this model had an area under the curve (AUC) of 0.90 (optimal odds ratio (OR), 44; 95 %CI 9-214) (Horgan et al. 2011).

The second most frequent targeted metabolite was homocysteine (Grandone et al. 2006; Choi et al. 2016), although levels were only differentiated between normal and SGA pregnancies when measured in 2nd trimester amniotic fluid, with a multiple linear regression model r^2=0.012 and p=0.029 (Grandone et al. 2006). Comparatively, the only common metabolite in 2nd trimester maternal hair was margarate, with conflicting results since it was found to be either increased (AUC 0.72, 95 %CI 0.58-0.86) (Delplancke et al. 2018) or decreased (Sulek et al. 2014). The N1,N12-diacetylspermine and the perfluorocarboxylic acids were associated to female SGA babies, not males. The former presented a 5-fold decreased risk of SGA across quintiles. The perfluorodecanoic and perfluoroundecanoic acids presented OR of 3.14 (95 %CI 1.07-9.19) and 1.83 (95 %CI 1.01-3.32) (Y. Wang et al. 2016). Tyrosine, an essential amino acid for infants, was part of the predictive model of maternal hair, combining 5 metabolites with an AUC of 0.998 (95 %CI 0.992-1.0) (Sulek et al. 2014). However, tyrosine did not predict SGA when urine samples were studied (Luthra et al. 2018). Methylmalonic acid (Choi et al. 2016), acetate, formate, or trimethylamine

(Luthra et al. 2018), did not differentiate SGA when compared to uncomplicated pregnancies ($p>0.05$).

Table 2.4: Predictive metabolites summarized according to their chemical class, subclass, and biological process

Predictive metabolites	Chemical class	Chemical subclass	Metabolic pathway
Margarate	Fatty acyls	Fatty acids and conjugates	Lipid transport, metabolism, peroxidation
Pentadecanoic acid	Fatty acyls	Fatty acids and conjugates	Lipid transport, metabolism, peroxidation; fatty acid metabolism and biosynthesis
Myristic acid	Fatty acyls	Fatty acids and conjugates	Lipid transport, metabolism, peroxidation; fatty acid metabolism and biosynthesis
Eicosatetraenoic acid	Fatty acyls	Fatty acids and conjugates	Lipid transport, metabolism, peroxidation; lipid metabolism pathway
Docosapentaenoic acid	Fatty acyls	Fatty acids and conjugates	Lipid transport and metabolism, fatty acid metabolism, alpha linolenic acid and linoleic acid metabolisms
Tyrosine [a]	Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Catecholamine biosynthesis; phenylalanine and tyrosine metabolism; thyroid hormone synthesis; transcription and translation
Homocysteine	Carboxylic acids and derivatives	Amino-acids, peptides, and analogues	Glycine and serine metabolism; methionine metabolism
Hexacosanedioic acid	Carboxylic acids and derivatives	Dicarboxylic acid and derivatives	Fatty acid biosynthesis
Sphinganine 1-phosphate	Sphingolipids	Phosphosphingolipids	Sphingolipid signalling pathway, nneuroactive ligand-receptor interaction
Sphingosine 1-phosphate	Sphingolipids	Phosphosphingolipids	Lipid metabolism pathway, sphingolipid metabolism
PFDeA	Alkyl halides	Alkyl fluorides	Not reported[b]
PFUnDA	Alkyl halides	Alkyl fluorides	Not reported[b]
25,OH,Vitamin D	Steroids and steroids derivatives	Vitamin D and derivatives	Lipid metabolism pathway
Diglyceride	Glycerolipids	Diradylglycerols	Adipocytokine signaling pathway
Lactate	Hydroxy acids and derivatives	Alpha hydroxy acids and derivatives	Gluconeogenesis, glycogenosis types IB and IC, pyruvate metabolism, triosephosphate isomerase
N1,N12-diacetylspermine	Carboximidic acids and derivatives	Carboximidic acids	
Lyso-phosphocholine	Glycerophospholipids	Glycerophosphocholines	Not reported[b]
2-methyloctadecanate	Saturated hydrocarbons	Alkanes	Not reported[b]
Levulinate	Keto acids and derivatives	Gamma-keto acids and derivatives	Not reported[b]

[a] Essential amino acid for infants. [b] No human metabolic pathways reported at KEGG. PFDeA: perfluorodecanoic acid; PFUnDA: perfluoroundecanoic acid.

2.4.5. Risk of bias and Applicability Concerns

Figure 2.2 shows synthesized data for all included studies. See **Appendix B** for individual QUADAS-2 data.

Regarding the risk of bias, all cohort studies conducted a consecutive participant inclusion (Delplancke et al. 2018; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; van Eijsden et al. 2008a). Nested case-controls matched cases and controls randomly](Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Gong et al. 2018) or according to maternal and infant characteristics (Horgan et al. 2011; Sulek et al. 2014; Luthra et al. 2018; Morillon et al. 2018). One study (Gong et al. 2018) failed to mention matching procedures ('Patient Selection' domain). Researchers were not blinded to SGA status when interpreting metabolomics results (Horgan et al. 2011; Sulek et al. 2014) (Delplancke et al. 2018; Ertl et al. 2012; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013; Gong et al. 2018), and thresholds of targeted metabolites were not pre-specified (Costet et al. 2011; Grandone et al. 2006; Y. Wang et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a) ('Index Test' domain). Conversely, SGA identification was not influenced by the metabolomics test, although it was unclear when laboratory experiments were performed in some studies (Sulek et al. 2014; Delplancke et al. 2018; Costet et al. 2011; Grandone et al. 2006; Choi et al. 2016; Gong et al. 2018). Birthweight charts were adequate, except for two studies. The first did not report which centile was chosen (Sulek et al. 2014), and the second used a centile designed for a different population (Grandone et al. 2006) ('Reference Test' domain). Two studies were ranked as 'high risk' because not all participants were included in the analysis (Costet et al. 2011; Ong et al. 2016) ('Flow and Timing' domain).

The QUADAS-2 tool also highlights the importance of how the findings of the included studies are suitable to the review question. In the Patient Selection domain, it was ranked as 'high applicability concerns' when infants born between the 4th and the 10^{th} centile, but with normal abdominal circumference growth velocity, were not included in final analysis (Gong et al. 2018). It was 'unclear' when the gestational age of maternal assessment was not standardized (Choi et al. 2016), or was inferred by

hair segment length (Delplancke et al. 2018); or when few metabolites from untargeted studies were chosen for interpretation (Gong et al. 2018) ('Index Test' domain). Finally, it was 'high' when the birthweight charts applied did not correspond to the study population (Sulek et al. 2014; Grandone et al. 2006) ('Reference Standard' domain).

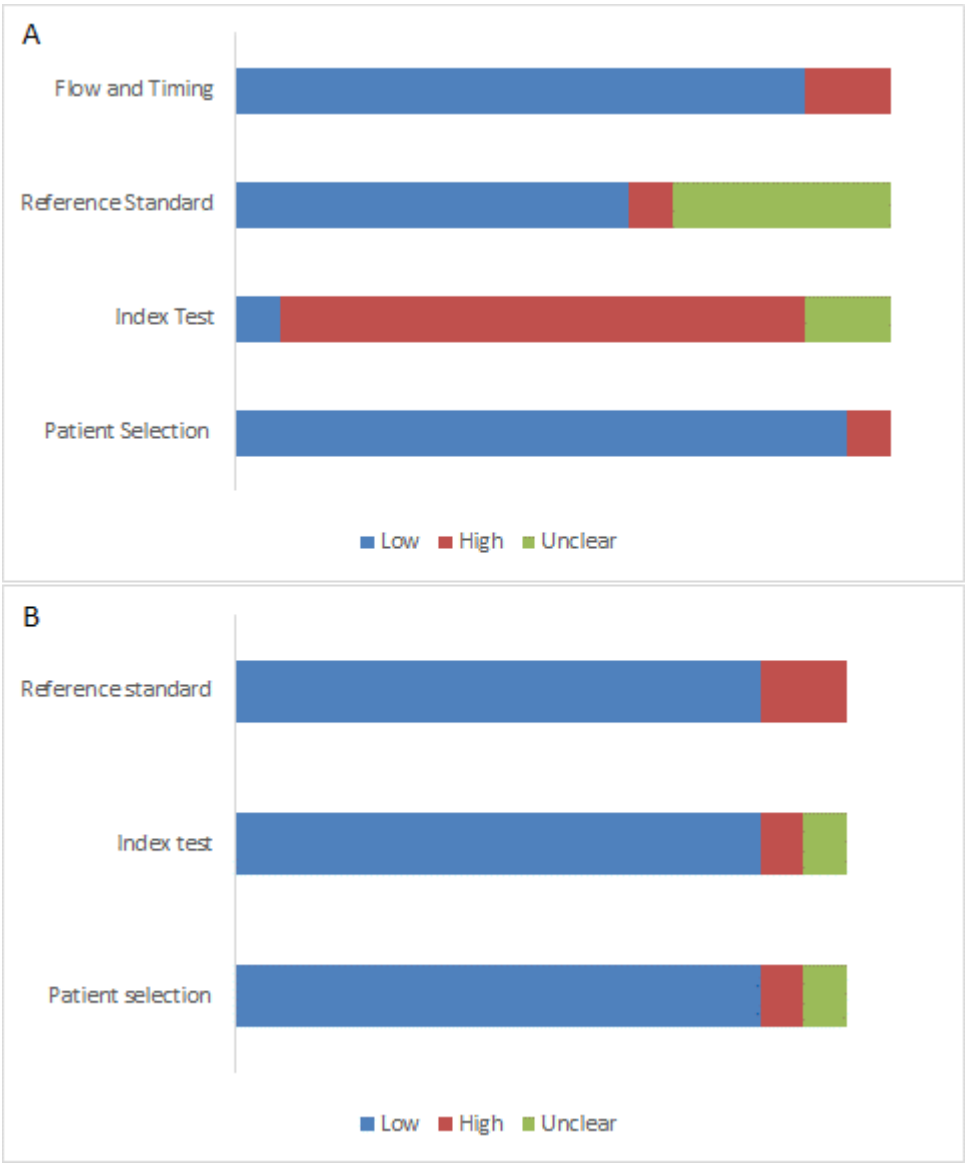


Figure 2.2: Assessment of risk of bias (A) and applicability concerns (B) of individual studies.

2.4.6. Meta-analysis

From the 15 included studies, only three were designed for prediction purposes (Horgan et al. 2011; Sulek et al. 2014; Morillon et al. 2018) and provided the AUC. The remaining reports described statistical differences of metabolites between SGA pregnancies and controls (Delplancke et al. 2018; Costet et al. 2011; Ertl et al. 2012; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013; Gong et al.

2018). Accuracy measures were extracted when available (**Table 2.3**). However, due to marked heterogeneity (**Tables 2.3** and **2.4)** of gestational age at sampling, type of samples used, type of birthweight chart chosen, thresholds for vitamin D deficiency, metabolomics approach, and identified compounds, a meta-analysis could not be performed.

2.5. Discussion

2.5.1. Main findings

In this first systematic review of metabolomics and adverse pregnancy endpoints, we presented techniques and metabolites, which were studied for the prediction of SGA. Any effect on birthweight has important implications for perinatal research, since it is related to short and long-term outcomes (N. Wang et al. 2017; Hales and Barker 2001; Melo et al. 2010; Ravelli et al. 1998), and in different generations (Chamorro-Garcia et al. 2017; Seferovic et al. 2015). Intrauterine environment influences fetal growth through epigenetic processes: altered gene expression potentially leads to distinct phenotypes (Padmanabhan et al. 2016). Metabolomics is the most adequate approach to study this outcome, since it is most directly related to phenotype (Dunn et al. 2011b).

Interpretation of metabolomics findings in pregnancy can be challenging. Firstly, maternal metabolites concentrations are influenced by placental transfer to and from the fetus. The 'mirror effect', seen for maternal plasma and venous cord blood metabolites at birth (Visentin et al. 2017), cannot be ruled out when only maternal specimens are studied. Secondly, maternal exposure to distinct compounds may affect metabolite levels. Statistically significant differences between SGA infants and controls may not express the totality of underlying pathological pathways and have no clinical meaning. Finally, it is unclear when the processes leading to SGA are initiated. The disruption in maternal metabolism can theoretically occur at any time. In general the lower the gestational age at which the condition is suspected, the more severe the phenotype will be at birth (Korzeniewski et al. 2017; Lees et al. 2013).

Thus, the description of clinical data in translational studies must deal with all these confounding factors.

Gestational age at sampling is probably the most important parameter for prediction purposes. With timely prediction, women could be referred to specialized care, have increased surveillance, and this in turn may lead to a reduction in perinatal mortality. There are temporal changes in the maternal metabolome during pregnancy (Delplancke et al. 2018; Luan et al. 2015; Di Giulio et al. 2004; Orczyk-Pawilowicz et al. 2016; Lindsay et al. 2015); therefore, it is reasonable to expect distinctive metabolites at different stages of pregnancy, as reported here. Unfortunately, a wide or unclear definition of gestational age of sampling (Choi et al. 2016; Y. Wang et al. 2016; Luthra et al. 2018; Gernand et al. 2013) render a more precise interpretation impossible, and may limit the clinical application of these results.

In contrast, gestational age at birth and birthweight centile seem to be the hallmarks of severity and prognosis of growth restriction (Mendez-Figueroa et al. 2016; Boghossian et al. 2018). Indeed, term and preterm SGA babies show distinct clinical phenotypes, and there are concerns that some babies <10th centile of birthweight are constitutionally small infants (Katz et al. 2014; Ray et al. 2009; Cheng et al. 2016). If only term deliveries are evaluated, the most severe cases of growth restriction may be potentially missed. Moreover, when term and preterm births are analyzed together, or when lower cutoffs are not specified (e.g. ≤3rd or ≤5th centile), the lack of predictive metabolites might mean that they are distinct conditions. Thus, we hypothesize that the predictive performance of metabolomics may be improved if data is analyzed by gestational age at delivery, and by different cutoffs of birthweight centiles.

Evidence suggests that tobacco smoke has an impact on birthweight (Abraham et al. 2017; Vesterinen et al. 2017; Dessì et al. 2018), although it is uncertain how and when fetal growth is impaired. It is possibly related to oxidative stress (Stone et al. 2014), and both maternal and fetal metabolism may be disturbed at delivery (Fischer et al. 2017; Rolle-Kampczyk et al. 2016). Studies that were included did not investigate cigarette-related chemicals or quantify exposure to tobacco smoke. Therefore, no relationship between SGA and tobacco was found. Hence, we suggest that tobacco

interferes with ongoing metabolic pathological processes, or its disturbance is related to additional metabolic pathways other than the one examined by the included studies.

2.5.2. Sub group findings

No reports have explored data on any maternal chronic condition, suspicion of SGA in pregnancy, or number of fetuses. The lack of clear statements about participant selection have hindered data interpretation and precluded these analyses.

The majority of included studies performed a targeted approach, i.e. a hypothesis-testing evaluation (Patti et al. 2012; Dunn et al. 2011b), driven by epidemiological or experimental data regarding SGA newborns. None of the targeted metabolites (Costet et al. 2011; Ertl et al. 2012; Grandone et al. 2006; Choi et al. 2016; Kiely et al. 2016; Y. Wang et al. 2016; Ong et al. 2016; Luthra et al. 2018; van Eijsden et al. 2008a; Gernand et al. 2013) were in common with those found by 'hypothesis-generating' metabolic profiling (Horgan et al. 2011; Sulek et al. 2014; Delplancke et al. 2018; Gong et al. 2018; Morillon et al. 2018) investigations. This reinforces the suggestion that various maternal metabolic pathways may be triggered by the SGA condition, and be detected by different biological samples. However, since blood is a very complex sample and GC-MS only evaluates volatile molecules (Dunn et al. 2011b), therefore our findings may be biased by study methodologies.

Untargeted studies, as expected, have characterized several metabolites that may be validated in future investigations. Nine lipids and fatty acid metabolites (Horgan et al. 2011; Sulek et al. 2014; Delplancke et al. 2018; van Eijsden et al. 2008a), two amino acids (Sulek et al. 2014; Grandone et al. 2006), and a steroid (Horgan et al. 2011; Ertl et al. 2012) have been identified as potential biomarkers of SGA.

All lipid-related metabolites identified are intermediates for energy storage and breakdown. Most metabolites were found in maternal blood (Horgan et al. 2011) or hair of the SGA group (Delplancke et al. 2018; Sulek et al. 2014). Blood levels of saturated and monounsaturated non-esterified fatty acids apparently remain stable throughout pregnancy, while long chain polyunsaturated fatty acid (DHA and EPA,

for example) measurements seem to show ethnicity-related changes (Lindsay et al. 2015). Experimental data shows the importance of hypoxia and oxidative stress to placental function and ultimately, to birthweight (Thomas et al. 2018; Määttä et al. 2018). Findings from included studies may represent a dysregulation of lipid pathways at the placental level, but an association with maternal background is unclear. Therefore, we hypothesize that disorders of lipid metabolism may be the 'metabolic snapshot' of defective deep placentation (Brosens et al. 2011), and might reflect maternal efforts to respond to impaired fetal growth.

Recommendations on the assessment of vitamin D and cutoffs to define vitamin D deficiency in pregnancy are controversial (Holick et al. 2011). However, vitamin D supplementation decreases SGA risk (Guang et al. 2018). In early pregnancy, vitamin D status has been related to SGA (Bodnar et al. 2010; Leffelaar et al. 2010), which is in accordance with this review, despite the inconsistent findings (Martínez-Domínguez et al. 2018). There is evidence that trophoblasts actively produce and secrete vitamin D metabolites (Park et al. 2017), but it is not clear how they mediate fetal growth impairment. Altered hepatic gene expression and liver function in vitamin D deficient female rats (S. S. Sharma et al. 2017), and single nucleotide polymorphisms (Barchitta et al. 2018) in vitamin D receptor gene have been suggested as mechanisms to be explored by a multidimensional omics approach.

Finally, homocysteine is an intermediate metabolite of the folate cycle. It is indirectly involved with DNA methylation and is a marker of folate deficiency (J. Kim et al. 2018). Maternal levels rarely reach hyperhomocysteinemia limits (Yajnik et al. 2014), but folate depletion (Smits and Essed 2001; van Eijsden et al. 2008b; Hogeveen et al. 2011) and homocysteine itself (Yajnik et al. 2014) are thought to be associated with a higher SGA risk. In this review, homocysteine was only statistically different in SGA pregnancies when measured in amniotic fluid (Grandone et al. 2006), although within the normal ranges proposed for 17-21 weeks (Imbard et al. 2013). Since amniocentesis is generally performed in women at higher obstetrical risk, future studies should investigate whether homocysteine in amniotic fluid represents a confounding factor or a new biomarker (Heazell et al. 2018).

2.5.3. Methodological quality

Most studies were ranked as 'low risk' of bias or applicability to the review question. However, the lack of clear descriptions of laboratory experiments, including sample preparation and storage, and blinding of the researchers to the case/control status, are major pitfalls of the included studies.

2.5.4. Strengths and limitations

To our knowledge, this is the first systematic review of metabolomics and an adverse pregnancy outcome (SGA). We presented possible biomarkers of SGA pathophysiology, metabolites implicated in lipid transport and metabolic pathways, as well as gluconeogenesis.

However, this analysis has some limitations. First, included studies showed heterogeneity, which is fundamental in systematic reviews. Indeed, there was a wide variety of participant characteristics and methods used, and not all authors provided a detailed description of methods employed. Although the Metabolomics Standard Initiative was released in 2007 (Goodacre et al. 2007), there is still poor adherence to guidelines (E. Considine et al. 2018; Spicer et al. 2017). Clear reporting (Xia et al. 2013; E. Considine et al. 2018; Spicer et al. 2017) and data sharing in repositories are crucial steps in identifying features of interest, specifically possible biomarkers to be validated in the clinical studies (Xia et al. 2013). Secondly, we could not perform a meta-analysis of the extracted data, impacting the translational potential of metabolomics.

Thirdly, we considered that birthweight was a surrogate measure of intrauterine development. SGA and FGR are not interchangeable concepts. However, SGA has been used as a surrogate for FGR in many clinical studies due to difficulties in defining optimal intrauterine growth: (i) FGR diagnosis relies mostly on ultrasound measurements of fetal biometry (Gordijn et al. 2016; Lesley M. McCowan et al. 2018), which in turn is subject to systematic errors (Lappen and Myers 2017); (ii) intrauterine development is adaptive, rather than uniform (Papageorghiou et al. 2014) or only genetically driven (Padmanabhan et al. 2016); (iii) growth impairment

at birth better identifies adverse neonatal outcomes than during pregnancy (Boghossian et al. 2018). It is recognized that changes in obstetric care occur when growth restriction is suspected, and neonatal outcomes are improved (Monier et al. 2016; Verlijsdonk et al. 2012). Thus, an accurate prediction of SGA during pregnancy will be a turning point in modern obstetrics.

2.6. Conclusions and implications for practice

Using the available clinical tools, efforts to predict SGA remain disappointing. Since SGA is a heterogeneous condition, it benefits from metabolomics. This novel area of research allows analysis of numerous types of biological fluids and detects thousands of metabolites in complex samples (Xia et al. 2013; Patti et al. 2012; David S. Wishart et al. 2018). However, findings of this systematic review must be interpreted with caution. The type of samples used may have influenced LC-MS (2^{nd} trimester maternal blood) and GC-MS (2^{nd} trimester maternal hair) findings in individual studies. Furthermore, the prediction of SGA in the context of maternal disorders, suspected FGR and twin pregnancies is an open field for future metabolomics studies, and environmental exposure investigation as well.

Surprisingly, none of the studies used ≤3^{rd} centile of birthweight as a cutoff or analyzed preterm deliveries and hypertensive syndromes. Considering our findings and the different phenotypic manifestations of SGA, we envision a better performance when (i) cutoffs other than the 10^{th} centile are tested; (ii) data on gestational age at sampling and at birth are standardized; and (iii) other pregnancy-related syndromes are considered, especially hypertension. Thus, future metabolomics results should advance in these critical points.

Finally, all detected biomarkers were related to lipid pathways and energy metabolism. We consider that research efforts to predict SGA should focus on compounds involved in these pathways, up to the 2^{nd} trimester of pregnancy.

Chapter 3:

Glycerophospholipid and Detoxification Pathways Associated with Small for Gestation Age Pathophysiology: Discovery Metabolomics Analysis in the SCOPE Cohort.

Morillon A-C, Leite DFB, et al. ''Glycerophospholipid and detoxification pathways associated with small for gestation age pathophysiology: discovery metabolomics analysis in the SCOPE cohort.'' Metabolomics, 17(1), 5. doi:10.1007/s11306-020-01740-9 (Morillon et al. 2021)

3.1. Abstract

Introduction:

Small for gestational age (SGA) may be associated with neonatal morbidity and mortality. Our understanding of the molecular pathways implicated is poor.

Objectives:

Our aim was to determine the metabolic pathways involved in the pathophysiology of SGA and examine their variation between maternal biofluid samples.

Methods:

Plasma (Cork) and urine (Cork, Auckland) samples were collected at 20 weeks of gestation from nulliparous low-risk pregnant women participating in the SCOPE study. Women who delivered an SGA infant (birthweight <10th percentile) were matched to controls (uncomplicated pregnancies). Metabolomics (urine) and lipidomics (plasma) analyses were performed using liquid chromatography mass spectrometry. Features were ranked based on FDR adjusted p-values from empirical Bayes analysis, and significant features putatively identified.

Results:

Lipidomics plasma analysis revealed that 22 out of the 33 significantly altered lipids annotated were glycerophospholipids; all were detected in higher levels in SGA. Metabolomic analysis identified reduced expression of metabolites associated with detoxification (D-Glucuronic acid, Estriol-16-glucuronide), nutrient absorption and transport (Sulfolithocholic acid) pathways. No common pathways were observed in the pathophysiology between plasma and urine.

Conclusions:

This study suggests higher levels of glycerophospholipids and lower levels of specific urine metabolites are implicated in the pathophysiology of SGA. Further research is needed to confirm these findings in independent samples.

3.2. Introduction

Small for gestational age (SGA) is defined as an infant born with a birthweight less than the 10th percentile when compared to population (weight according to gestational age at birth) or to customised curves (charts individualized on the basis of maternal characteristics).(D. Sharma et al. 2016b) SGA is associated with placental dysfunction(Dessì et al. 2015) and is a major cause of maternal and neonatal morbidity.(Diderholm 2009) Infants born with SGA are at higher risks of perinatal mortality and neonatal complications, including asphyxia(Rosenberg 2008) and neurological impairment,(Grantham-McGregor 1998; D. Sharma et al. 2016a) as well as longer term morbidity, including increased risk of cardiovascular disorders and type 2 diabetes.(Dessì et al. 2012) The lower the birth centile, the higher the risk of short and long term morbidity. SGA can be further classified as moderate (birthweight centile in the 3rd to 10th percentile), and severe (birthweight less than the 3rd percentile).(P. A. Lee et al. 2003) A better understanding of the molecular pathways involved in the pathophysiology of SGA could enable better prevention, earlier detection and potential treatment of this pregnancy complication.

Metabolomics is the study of all small weight molecules (50-2000 Da), or metabolites, present in a sample. Metabolic profiling in a clinical research setting has led to the determination of risk factors and pathophysiology of specific diseases.(Dunn et al. 2011b) Three metabolomics studies using plasma or serum samples have shown that the physiopathology of SGA appears to affect lipid and fatty acid metabolism.(Horgan et al. 2011; Sulek et al. 2014; Delplancke et al. 2018) In the Greek Rhea cohort study, urine samples taken at 11 weeks of gestation, showed an association between elevated levels of six metabolites and higher risks of subsequent SGA and preterm birth.(Maitre et al. 2014)

We hypothesised that maternal urine and plasma would share common biological pathways that may explain pathophysiological changes seen in SGA and that these changes would precede detection of disease. The aim of our study was to gain further insight into the metabolic and lipidomic pathways involved in the pathophysiology of SGA, using urine and plasma metabolic profiles of women at 20 weeks of gestation, in geographically distinct populations of the SCOPE pregnancy cohort.

3.3. Materials and Methods

3.3.1. Participants

The present nested case-control study was performed on samples selected from SCreening fOr Pregnancy Endpoints study (SCOPE, www.scopestudy.net) in Cork, Ireland, and Auckland, New Zealand. This study was performed in accordance with the 1964 Helsinki declaration and its later amendments. Informed consent and ethical approval were obtained (Ireland ECM5 (10) 05/02/08, and New Zealand AKX/02/00/364). SCOPE is an international pregnancy cohort that recruited 1,773 women in Ireland and 2034 in New Zealand with low-risk and singleton pregnancies (L. C. Kenny et al. 2014). Selected participants who delivered small for gestational age babies, with customised birthweights less than the 10^{th} centile (cases), were matched to participants who had healthy and uncomplicated pregnancies (controls). Controls were matched to cases by age (± 5 years), body mass index (BMI ±3.5 kg/m^2), and ethnicity (**Table 3.1**). For definitions of all clinical endpoints used in **Table 3.1**, see Australian New Zealand Clinical Trials Registry (ANZCTR), using study number ACTRN12607000551493. Urine and plasma samples were obtained from 40 selected SGA cases in Cork population, and matched with 40 controls. In the Auckland population, only urine samples were analysed, with a similar case-control design: 40 women were selected for the SGA group and matched with 40 controls. The nested case-control study designs are summarised in **Figure. 3.1**.

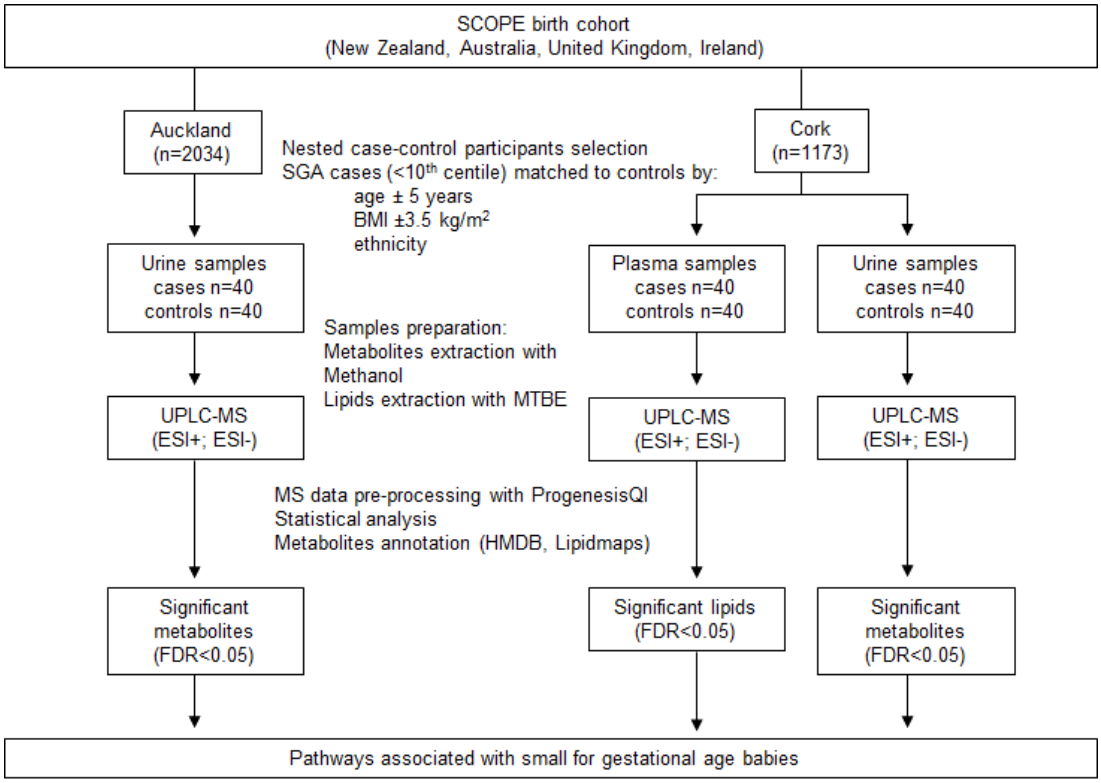


Figure 3.1: Description of selection of participants, plasma and urine samples preparation and analysis.

3.3.2. Reagents and materials

Liquid chromatography (LC) grade methanol, methyl tert-butyl ether (MTBE), acetonitrile (ACN), isopropanol (IPA), formic acid, ammonium formate, and glass tubes were obtained from Fisher Scientific (Loughborough, UK). LC-MS glass vials were purchased from Waters (Waters, Dublin, Ireland).

3.3.3. Sample preparations

3.3.3.1. Plasma samples

Plasma samples taken at 20 weeks of gestation were prepared in a randomised order. The lipid extraction method used was based on a protocol described by Matyash et al.(Matyash et al. 2008) Plasma samples were taken out of -80°C storage and left on ice until thawed; none had been previously thawed. Plasma samples (200 µl) were transferred to glass tubes, pre-chilled methanol (-20 °C, 600 µl) was added, and

vortex mixed for 1 minute. MTBE was added (5 ml) to the samples, which were then incubated at room temperature on a shaker for 1 hour. Water (1 ml) was added and samples were left on a shaker for 10 minutes at room temperature; samples were then centrifuged at room temperature for 10 minutes at 1000 g. The organic (top) layer was transferred into glass tubes and left to evaporate overnight, at room temperature. Each sample was reconstituted in 200 µl of 65:30:5, IPA:ACN:Water and vortex mixed for 30 seconds. Lastly, all samples were transferred to LC glass vials in preparation for LC-MS analysis. Quality control (QC) samples were created by pooling 10 µl from each extracted sample. Multiple aliquots of pooled QC (100 µl per aliquot) were constructed using several LC vials.

3.3.3.2. Urine samples

Urine samples taken at 20 weeks of gestation were prepared in a randomised order, following the Nature Protocol developed by Want *et al.*(Want et al. 2010) Samples stored at -80°C, were defrosted on ice, then centrifuged for 10 minutes at 10,000 *g*. Supernatant (100 µl) was transferred to a new tube and milliQ water with 0.1 % of formic acid (FA, 200 µl) was added (ratio 1:2, urine: water 0.1 %FA). Samples were then transferred to LC-MS glass vials. Quality control samples (QCs) were prepared by pooling 20 µl from each sample; sample preparation was then the same as for other samples with 100 µl of pooled QC subsequently transferred to several LC vials.

3.3.4. LC analysis

3.3.4.1. Plasma samples

LC-MS analysis was performed using an ACQUITY ultra-high performance liquid chromatography (Waters Corp, Milford, MA) coupled with a Synapt G2-S quadrupole time-of-flight (UPLC-Q-TOF) mass spectrometer (Waters Corp, Wilmslow, UK). Samples were analysed on a BEH C18 column (1.7 µm, 2.1 x 100 mm) which was maintained at 65 °C, whilst samples were maintained at 7 °C. Analytes were separated over a 23 minute gradient using a flow rate of 0.4 mL/min. Mobile phase A was a mix of 10 mM of ammonium formate in acetonitrile: water (ACN:H_2O, 60:40

(v:v)), and B was a mix of 10 mM of ammonium formate in isopropanol: acetonitrile (IPA: ACN, 90:10 (v:v)). The gradient consisted of initial conditions at 30 % of B, before increasing to 99 % of B at 18 minutes and maintaining this for a further 2 minutes before decreasing to 30 % of B over 2 minutes, returning to initial conditions. Column conditioning consisted of 8 repeat injections of the pooled QC. Samples were analysed as technical triplicates in a randomised order with the pooled QC injected every tenth injection throughout the analysis.

3.3.4.2. Urine samples

LC-MS analysis was performed using an ACQUITY ultra-high performance liquid chromatography (Waters Corp, Milford, MA) coupled with a Synapt G2-S quadrupole time-of-flight (UPLC-Q-TOF) mass spectrometer (Waters Corp, Wilmslow, UK). Samples were analysed on a BEH C18 column (1.7 µm, 2.1 x 100 mm) which was maintained at 40 °C, whilst samples were maintained at 7 °C. Analytes were separated over a 15 minutes gradient using a flow rate of 0.5 mL/min. Mobile phase A was water and 0.1 % formic acid, and B was methanol and 0.1 % formic acid. The gradient consisted of initial conditions at 1 % of B for 1 minute, before increasing to 15 % of B over 3 minutes, increasing to 50 % of B over 3 minutes, and further increasing to 95 % of B over 3 minutes and maintaining this for a further 1 minute, before decreasing to 1 % of B over 5 minutes, returning to initial conditions.

Column conditioning consisted of 8 repeat injections of the pooled QC. Samples were analysed as technical triplicates in a randomised order with the pooled QC injected every tenth injection throughout the analysis.

Cork samples were analysed first, followed by Auckland samples analysis at a different date and on a different BEH C18 column, following the protocols described above.

3.3.5. MS configuration

For both urine and plasma samples, data were acquired using the data independent acquisition (DIA) mode, MS^E (Bateman et al. 2002; Silva et al. 2005) using a Synapt G2-S Q-TOF mass spectrometer (Waters Corp, Wilmslow, UK). Data were acquired in resolution mode, from 50 to 1500 Da, and in both positive and negative electrospray ionisation modes (ESI+, and ESI-). Both precursor (low energy) and fragment (high energy) ion data were collected during the same acquisition, with 0.1 second scan time for each, and a total cycle time of 0.2 seconds. A linear collision energy ramp (20-40 eV) was applied for high energy, over the 0.1 second scan. Capillary voltage was set at 3.0 kV, sampling cone at 40 V, extraction cone to 5 V. The source temperature was set at 120 °C, and the desolvation temperature was 650 °C. The desolvation gas flow rate was set at 800 L/h, and cone gas at 50 L/h. Mass calibration was performed using a sodium formate mix (Waters, Wexford, Ireland), recommended by the manufacturer before each batch analysis. Real time lock mass correction was performed using a leucine enkephalin (LeuEnk, 1ng/µL) mix, injected at 10 µL/min through a lock-spray probe and acquired every 30 seconds.

3.3.6. Data processing

Using Progenesis QI version 2.4 (Nonlinear dynamics, Newcastle, UK) was used to align and peak pick the LC-MS data. An appropriate pooled QC was selected as the alignment reference to chromatographically align the data. Data were then peak picked, and considered the adducts corresponding with M+H, M+H-$2H_2O$, M+NH_4, M+Na, M+K, M+2H, M+2Na (ESI+), or M-H, M-H_2O-H, M-Na-2H, M+K-2H, M+Cl, M-2H (ESI-). Downstream statistical analysis was performed using the compound measurements exported from Progenesis QI.

3.3.7. statistical analysis

Statistical analysis of the demographics and clinical data was performed using Student T test, Mann-Whitney U test, or Pearson χ^2 test, with multiple testing corrections as appropriate (IBM SPSS Statistics 24). Results were considered

statistically significant when p-values were <0.05 (**Table 3.1**). Before sample preparation and analysis, a block randomisation based on the patients' BMI and the outcome was performed. No significant dependency between measurement order, the outcome and biometric and clinical information about patients was observed using Mann Whitney U test, Spearman correlation, Chi square test and Kruskal-Wallis test as applicable; Benjamini and Hochberg step-up False Discovery Rate-controlling procedure was applied for multiple testing correction.(Benjamini and Hochberg 1995)

Statistical analyses of the UPLC-MS data were performed using the R statistical software(R Core Team 2013) and the Bioconductor package limma(Ritchie et al. 2015), and package ggplot2(Wickham 2011) was used to create volcano plots. Data obtained from analyses of urine and plasma samples were subjected to the same statistical analyses performed independently. Data were acquired in positive and negative electrospray ionisation modes (ESI+, ESI-), which are known to show significant differences, and were therefore analysed independently. Median normalisation of the raw UPLC-MS data, and application of quality control procedures(Broadhurst et al. 2018) were performed. Measurement precision of each feature was checked by computing coefficient of variation (CV) and the missing rate over the replicate measurements. Features with a missing rate greater or equal to 20 % and features with a CV greater or equal to 30 % were filtered out. Robust multi-variable methods(F. Wang et al. 2018; Brereton and Lloyd 2014) were used to rank and select features. Empirical Bayes was used on each feature, and the design was adjusted for replicate measures(Smyth 2004); features analysed were ranked by adjusted p-value. In addition, the Mann Whitney U test was performed on the average measurement per patient with multiple testing correction.(Benjamini and Hochberg 1995) Agreement between the two methods was shown and confirmed the robustness of the results (Data not shown). The features of interest were tested for correlation with clinical variables, using Wilcoxon-Mann-Whitney Test, Spearman Correlation Test, or Kruskal-Wallis Test as appropriate. Multiple testing correction(Benjamini and Hochberg 1995), with a false discovery rate (FDR) cut-off of 25 %, was applied.

Table 3.1: Characteristics of participants and pregnancy outcomes in Auckland and Cork populations.

Variables	Auckland			Cork		
	Control (n=40)	Case (n=40)	p-value	Control (n=40)	Case (n=40)	p-value
Participants characteristics:						
Ethnicity						
Caucasian	35 (87.5%)	35 (87.5%)	1.0	40 (100%)	40 (100%)	NA
Other [a]	5 (12.5%)	5 (12.5%)		0	0	
Socioeconomics index	50 (36.75-50.75)	48 (39.25-59.75)	0.615	43.5 (29.0-50.0)	30 (22.00-45.75)	0.08
Age (years)	31.00 (28.00-32.75)	31.00 (29.00-34.00)	0.448	29 (23.5-31)	28 (24.25-31)	0.412
Primigravida	31 (77.5%)	24 (60%)	0.147 [c]	35 (87.5%)	33 (82.5%)	0.755 [c]
BMI (kg/m^2) at 15weeks of gestation	23.35 (22.33-26.08)	24.30 (21.68-27.20)	0.969	24.65 (21.58-27.83)	24.85 (23.13-29.38)	0.607
At 20 weeks of gestation:						
Current smoker	3 (7.5%)	2 (5%)	1.0 [c]	4 (10%)	12 (30%)	0.048 [c]
Current alcohol consumer	7 (17.5%)	5 (12.5%)	0.755 [c]	6 (15%)	6 (15%)	1.0 [c]
Current drug user	2 (5%)	1 (2.5%)	1.0 [c]	0	0	NA
Gestation (weeks)	19.86 (19.18-20.4)	20.14 (19.43-20.43)	0.228	20.86 (20.33-21.14)	20.71 (20.04-21.14)	0.396
Pregnancy outcome:						
GDM						
Yes	0	1 (2.5%)	0.574	0	0	0.31 [c]
No	35 (87.5%)	35 (87.5%)		8 (20%)	13 (32.5%)	
Unknown	5 (12.5%)	4 (10%)		32 (80%)	27 (67.5%)	
Diagnosed with PE	0	0	NA	0	0	NA
Diagnosed with GH	0	0	NA	0	1 (2.5%)	1.0 [c]
sPTB, < 37w	0	1 (2.5%)	1.0 [c]	0	0	NA
Gestational age at delivery (weeks)	40.22 (39.14-41.07)	39.57 (38.9-40.43)	0.214	40.64 (39.9-41.25)	39.86 (38.25-41)	0.014
Customised birthweight centile [b]	54 (30.25-78.75)	3.4 (2-5)	<0.001	49 (35.45-65.83)	2.35 (0.8-4.48)	<0.001

Values are shown as mean (SD), median (interquartile range) or n (%). a including African, Asian, Maori and Pacific Islander; b adjusted for mother's height, weight at 15 weeks visit, ethnicity, sex and weight of baby and gestation at delivery of baby; c adjusted with Fisher's Exact Test; NA: Not applicable; BMI: Body mass index; GDM: gestational diabetes mellitus; PE: pre-eclampsia; GH: gestational hypertension; sPTB: spontaneous preterm birth; all mothers were nulliparous.

3.3.8. Metabolites putative annotation

In accordance with the MSI reporting standards, we have achieved metabolite identification level 2, or putatively annotated compounds.(Salek et al. 2013). For each dataset, the exact mass of significant features (adjusted p-value < 0.05) were searched against Human metabolome database (HMDB, version 4.0)(David S. Wishart et al. 2018) and Lipidmaps (version of January 2019)(Cotter et al. 2007) using Progenesis QI identification tool. The search parameters were set for an exact mass tolerance of 5 ppm for the precursor ion, and of 10 ppm for the fragment ion. With UPLC-MS, it is common that metabolites are detected multiple times due to fragmentation, dimerization, chemical adduction, or multiple charging. Putative annotations were reported for unique metabolites, after removing duplicates and metabolites of drugs or metabolites originating from food, and after checking the retention time. Using HMDB, the metabolites reported were grouped into chemical classes.

3.4. Results

3.4.1. Global lipidomic analysis of plasma samples

Clinical and demographics data from selected Cork SGA cases (customised birthweight <10th centile, and median 2.35, IQR 0.8-4.48, n=40) and matched controls are presented in **Table 3.1**. Cases were matched to controls by ethnicity, BMI and age and no significant differences were observed in these parameters. Smoking status was significantly different between cases and controls, as was gestational age at delivery (mean 39.19 (SD 2.84) vs 40.41 (SD 0.94) weeks gestation; p =0.007).

In plasma samples, 8,328 (ESI+) and 6,842 (ESI-) features were detected. After normalisation and data filtration, 6,179 (ESI+) and 5,048 (ESI-) features were further analysed, and 1,167 (ESI+) and 788 (ESI-) features showed adjusted p<0.05. Thirty-three features (22 in ESI+ and 12 in ESI-) were matched to known lipids and are reported (**Table 3.2**). Among the 33 lipids putatively annotated, 22 were glycerophospholipids (GPL), including 3 phosphatidylethanolamines (PE), 5 phosphatidylserines (PS), 3 phosphatidylcholines (PC) and 1 lysoPC, 1

phosphatidylglycerophosphate (PA) and 1 lysoPA, 2 phosphatidylinositols (PI), 2 phosphatidylglycerophosphates (PGP), 3 phosphatidylglycerols (PG), and 1 cardiolipin (CL). With the exception of CL(72:2), all of these glycerophospholipids were detected at higher levels in cases when compared to controls. In addition to the glycerophospholipids, 4 sphingolipids were putatively annotated: 2 ceramides (Cer), 1 ganglioside, and 1 sphingomyeline (SM). These 4 sphingolipids were detected in higher levels in cases compared to controls. Finally, 3 fatty acyls, 3 glycerolipids, one organooxygen compound, and one cholesterol fatty ester were identified in Cork plasma samples (**Table 3.1, Figure 3.2**). Box plots showing the distribution of the lipids of interest in controls (n=40), cases with a customised birthweight centile below the 3rd (n=22), and cases with a customised birthweight between the 3rd and 10th centile (n=18) are presented in **Figure 3.3**.

In addition, we determined whether there was a correlation between the lipids of interest and either smoking status, or gestational age at delivery. None of the lipids of interest correlated significantly with gestational age at delivery. However, one lipid was significantly associated with maternal smoking status: CL(72:2) (p-value = 5.82×10^{-4} with FDR = 0.06) and was excluded from our results. Box plots showing the levels of this lipid in smokers and non-smokers are shown in **Figure 3.4**.

Table 3.2: Lipids that showed significant (adjusted p-value <0.05) differences in SGA group compared to controls in Cork plasma samples.

Lipids	Chemical class	Adjusted p-value	Fold change	Direction in SGA group
PE(P-31:0)	Glycerophospholipids	2.89E-03	1.17	UP
PE(42:1)	Glycerophospholipids	1.23E-04	1.36	UP
PE(36:4)	Glycerophospholipids	9.68E-04	1.18	UP
PS(O-37:0)	Glycerophospholipids	2.55E-06	1.39	UP
PS(41:5)	Glycerophospholipids	1.60E-04	1.36	UP
PS(37:2)	Glycerophospholipids	1.62E-04	1.29	UP
PS(43:6)	Glycerophospholipids	4.64E-04	1.54	UP
PS(P-34:0)	Glycerophospholipids	1.37E-02	1.22	UP
PC(O-42:4)	Glycerophospholipids	1.24E-02	1.15	UP
PC(40:5)	Glycerophospholipids	9.68E-04	1.25	UP
PC(38:6)	Glycerophospholipids	6.68E-05	1.25	UP
LysoPC(16:0)	Glycerophospholipids	8.57E-04	1.15	UP
PA(O-36:2)	Glycerophospholipids	3.17E-02	1.16	UP
LysoPA(18:1)	Glycerophospholipids	7.53E-04	1.15	UP
PI(37:1)	Glycerophospholipids	1.99E-04	1.20	UP
PI(P-33:1)	Glycerophospholipids	3.51E-04	1.23	UP
PGP(38:4)	Glycerophospholipids	1.13E-05	1.35	UP
PGP(40:4)	Glycerophospholipids	4.87E-05	1.35	UP
PG(36:6)	Glycerophospholipids	5.94E-04	1.23	UP
PG(39:8)	Glycerophospholipids	6.25E-04	1.29	UP
PG(38:4)	Glycerophospholipids	6.43E-04	1.20	UP
CL(72:2)	Glycerophospholipids	9.75E-04	1.31	DOWN
DG(44:4)	Glycerolipids	4.25E-02	1.12	UP
DG(O-34:1)	Glycerolipids	1.73E-02	1.36	UP
TG(64:15)	Glycerolipids	4.84E-04	2.15	DOWN
N,N-dimethyl arachidonoyl amine	Fatty acyl	6.69E-19	7.86	UP
N-palmitoyl valine	Fatty acyl	8.58E-11	1.99	UP
8S-hydroxy-hexadecanoic acid	Fatty acyl	4.93E-04	1.88	DOWN
SM(34:1)	Sphingolipids	6.30E-05	1.25	UP
Ganglioside GA2 (40:1)	Sphingolipids	2.77E-02	1.56	UP
Cer(34:0)	Sphingolipids	3.30E-02	1.26	UP
Cer(39:2)	Sphingolipids	3.77E-02	1.16	UP
CE(17:0)	Steroids and steroid derivatives	3.01E-04	1.28	DOWN

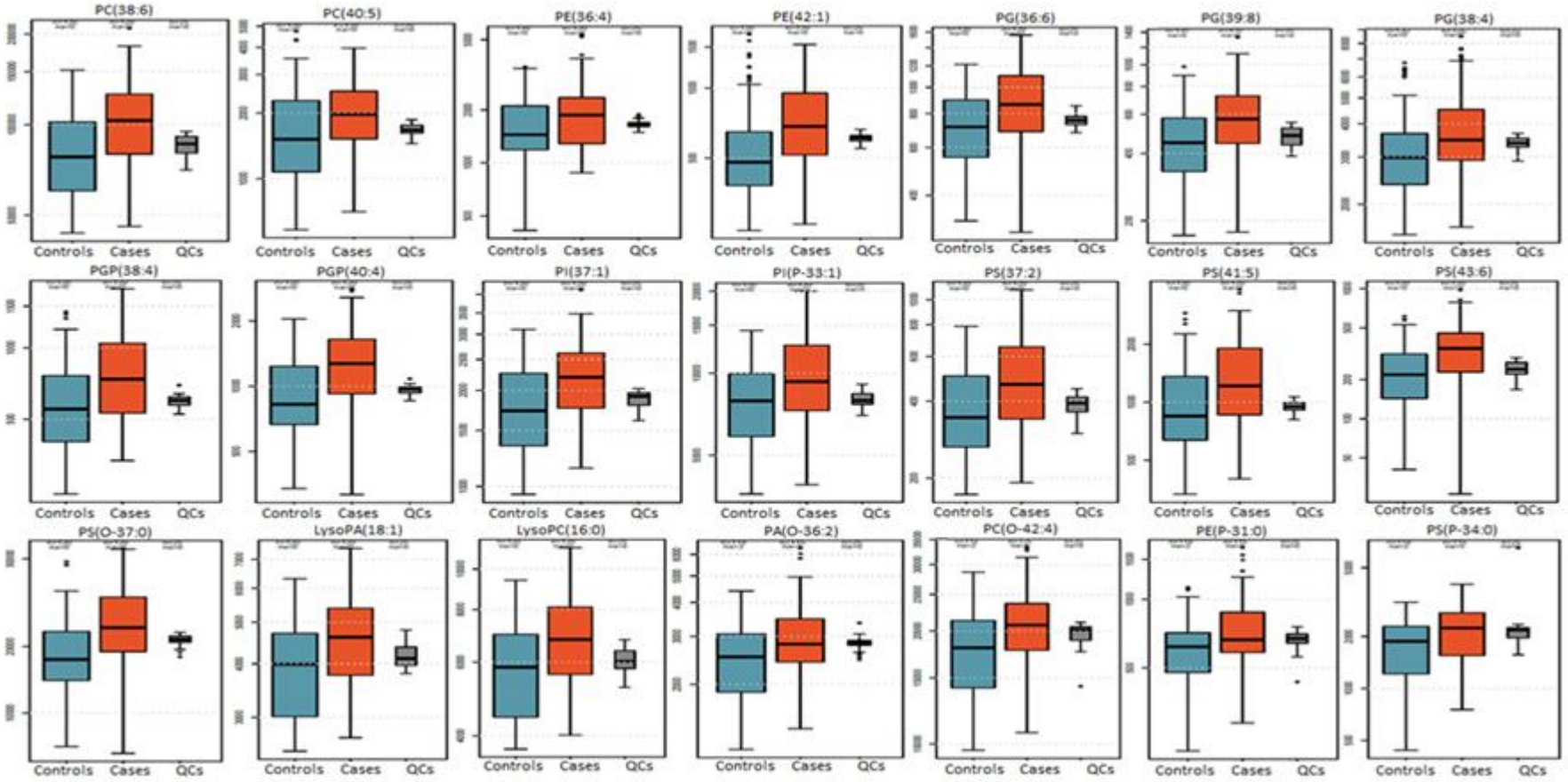


Figure 3.2: Box plots showing the normalised intensity of features significantly altered (adjusted p-value < 0.05) in small for gestational age group (cases, orange box) compared to control group (blue box), from untargeted UPLC-MS analysis of SCOPE Cork plasma samples (cases n=40, controls n=40). Quality control (QCs) group also shown (grey box).

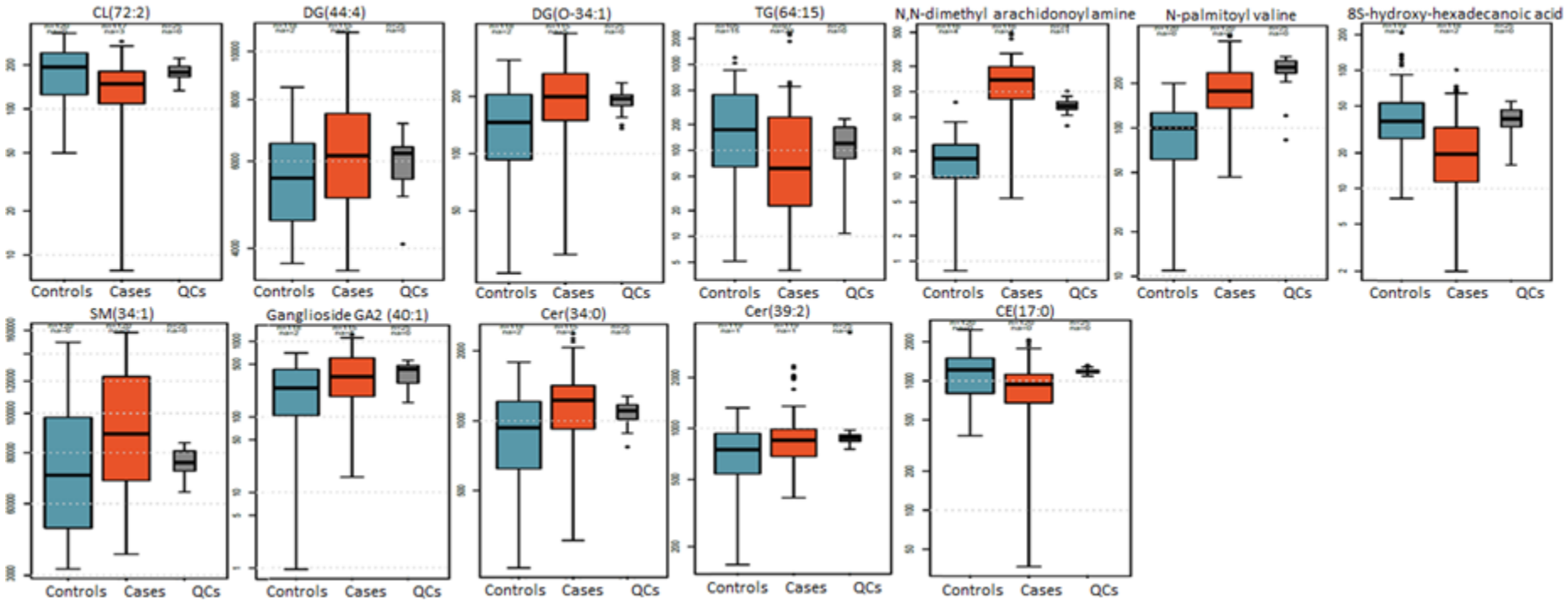


Figure 3.2: Continued**.**

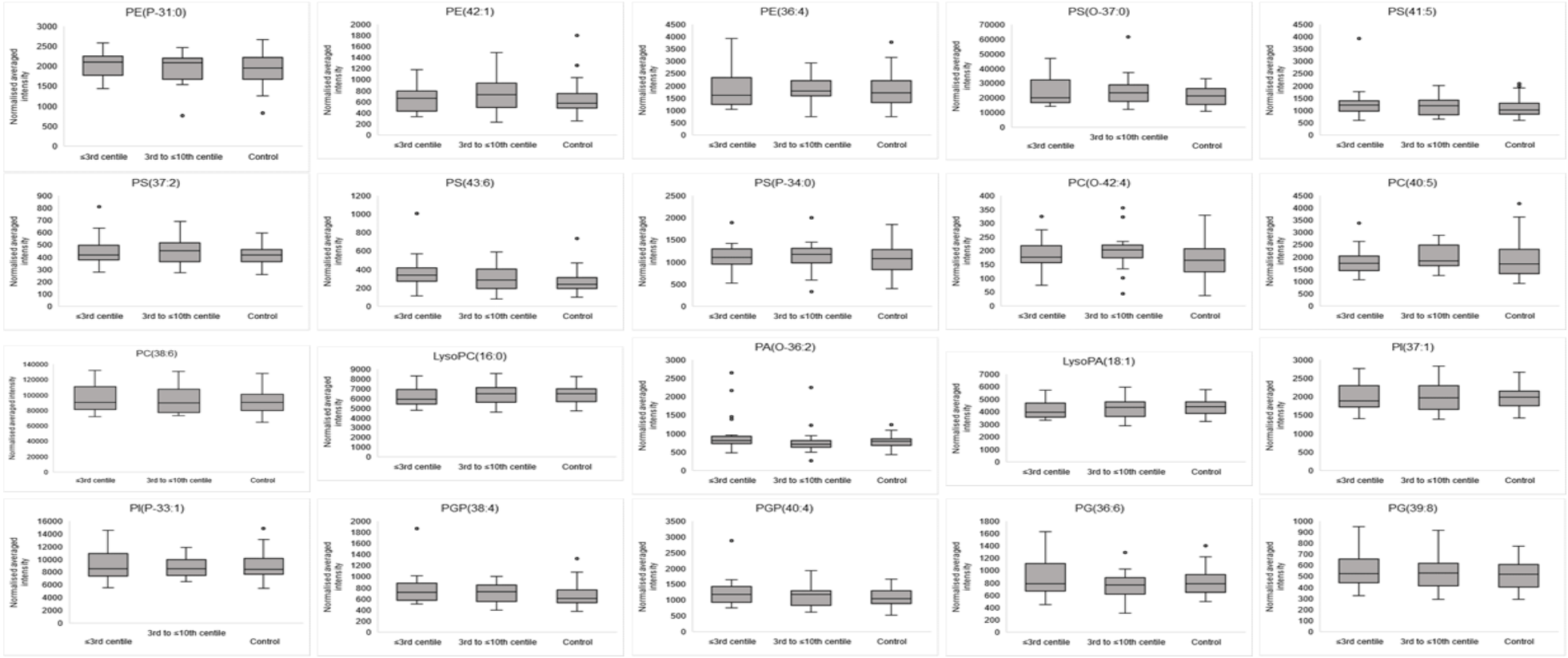


Figure 3.3: Box plot showing the normalised and averaged intensity of lipids of interest in three groups from untargeted UPLC-MS analysis of SCOPE Cork plasma samples. Cases are small for gestational age (SGA) with a customised birthweight below the 3rd centile (n=22), or between the 3rd and 10th centile (n18), and controls (n=40).

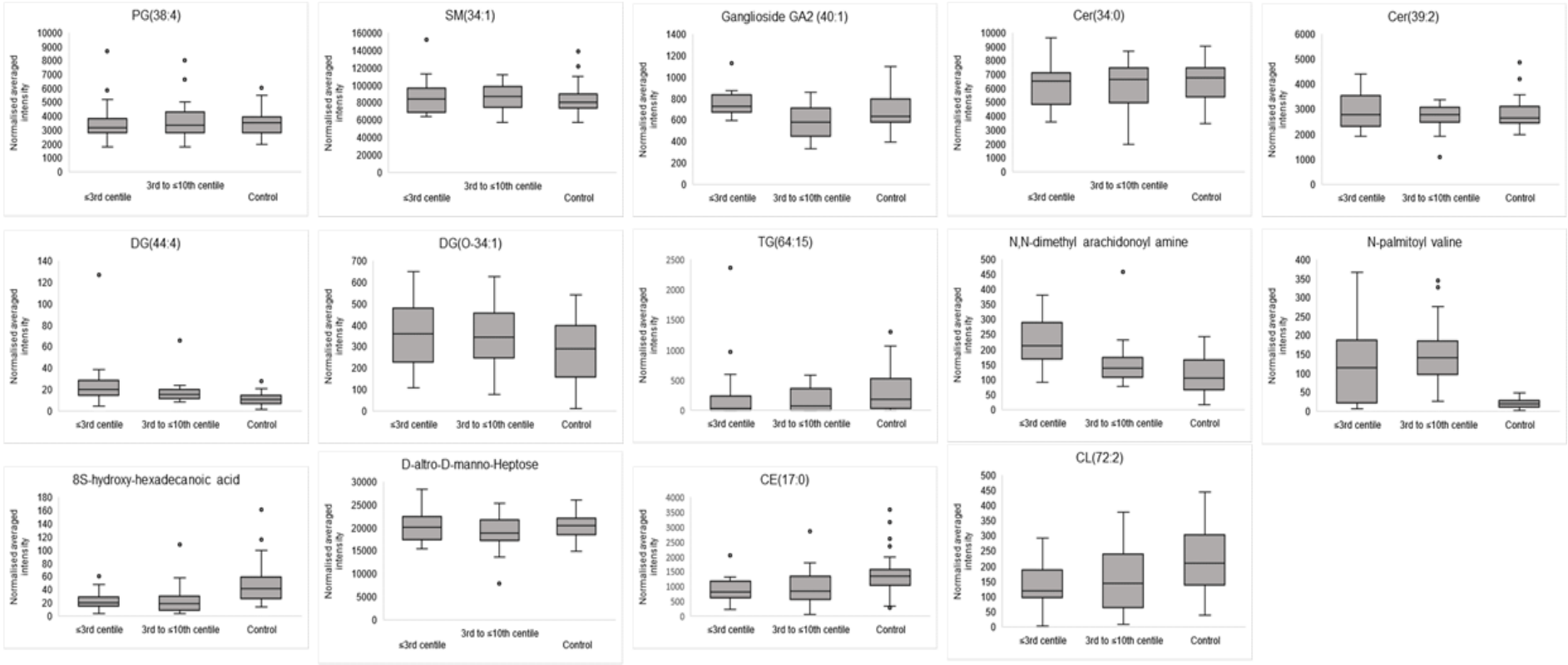


Figure 3.3: Continued.

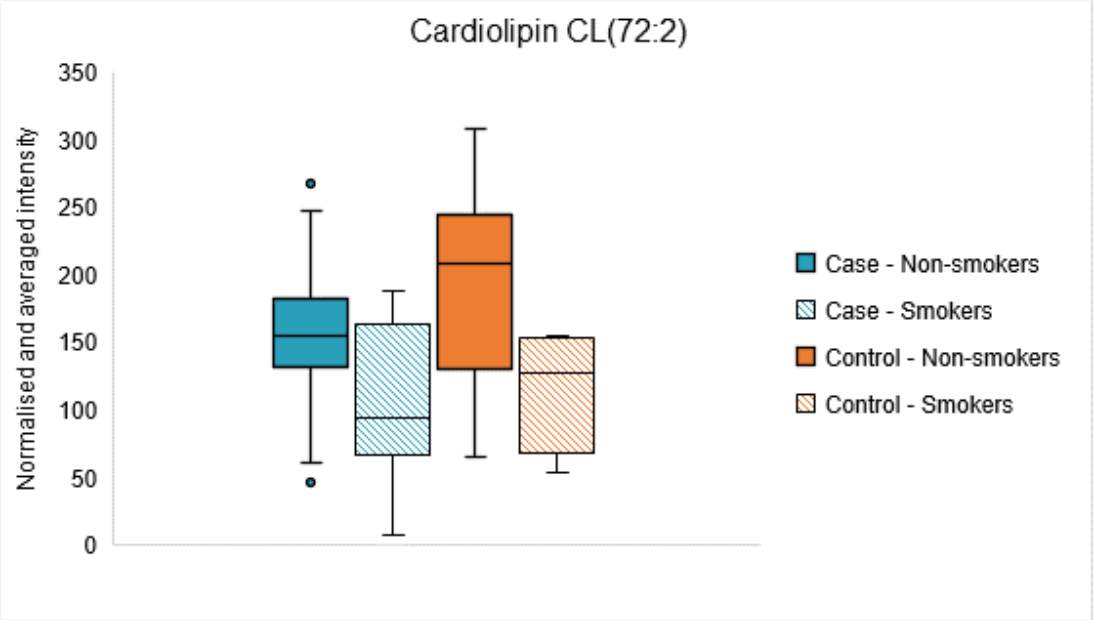


Figure 3.4: Box plot showing the normalised and averaged intensity of cardiolipin CL(72:2) in four groups from untargeted UPLC-MS analysis of SCOPE Cork plasma samples. Cases are small for gestational age (SGA, customised birthweight ≤10th centile). Blue box: case non- smokers (n=28), Blue pattern box: case smokers (n=12), Orange box: control non-smoker (n=36), and Orange pattern box: control smokers (n=4).

3.4.2. Global metabolomics analysis of urine samples

Clinical data and demographics of selected women are described in **Table 3.1**. Importantly, no significant correlations were found between the metabolites of interest and either smoking status, or gestational age at delivery, in either the Cork or Auckland populations.

In Auckland, urine metabolite profiles detected 11,488 (ESI+) and 10,513 (ESI-) features. After data normalisation and filtration, 11,213 (ESI+) and 10,216 (ESI-) features were further analysed, and 681 (ESI+) and 948 (ESI-) features showed significant differences between cases and controls (adjusted p <0.05). After a database search, as described in methods section, 3 features (2 in ESI+, 1 in ESI-) matching with known metabolites are reported (**Table 3.3**): 1 glucuronic acid (D-glucuronic acid), 1 organooxygen compound (beta-1,4-Mannosyl-N-

acetylglucosamine), and 1 steroid derivative (18-Hydroxycortisol). These metabolites were detected in lower levels in cases compared to controls.

In Cork, urine metabolite profiles detected 16,008 (ESI+) and 8,853 (ESI-) features. After data normalisation and filtration by CV and missing rate, 15,468 (ESI+) and 8,838 (ESI-) features were further analysed, and 271 (ESI+) and 866 (ESI-) features showed significant differences between cases and controls (adjusted $p < 0.05$). After a database search, as described in methods section, 4 features (1 in ESI+, 3 in ESI-) matching with known metabolites are reported (**Table 3.3**): 1 bile acid (sulfolithocholic acid), 1 steroid derivative (estriol-16-Glucuronide), 1 benzaldehyde (4-Hydroxybenzaldehyde), and 1 neuropeptide (Neuromedin N (1-4)). All four were detected in lower levels in cases compared to controls.

Plasma and urine samples from the same Cork participants were analysed. However, no common metabolite features were identified in both urine and plasma.

Table 3.3: Significant (adjusted p-value <0.05) metabolites of interest detected in urine samples

Metabolite	Chemical class	Biological function or pathway	Adjusted p-value	Fold change	Direction in SGA group	Population
Sulfolithocholic acid	Bile acid	Nutrient absorption and transport Lipid metabolism pathway	7.66E-03	1.12	Down	Cork
Estriol-16-Glucuronide	Steroid and steroid derivatives	Detoxification Lipid metabolism pathway	3.83E-04	1.12	Down	Cork
4-Hydroxybenzaldehyde	Hydroxybenzaldehyde		4.79E-02	1.17	Down	Cork
Neuromedin N (1-4)	Carboxylic acids and derivatives, neuropeptide		2.71E-02	1.02	Down	Cork
D-Glucuronic acid	Glucuronic acid	Detoxification	4.10E-02	1.05	Down	Auckland
18-Hydroxycortisol	Steroid and steroid derivatives	Lipid metabolism pathway	4.36E-02	1.13	Down	Auckland
beta-1,4-Mannosyl-N-acetylglucosamine	Organooxygen compounds		3.88E-02	1.05	Down	Auckland

3.5. Discussion and conclusions

This study has highlighted the following main findings: i) altered glycerophospholipids and sphingolipids in plasma samples from women with SGA pregnancies, and ii) lower levels of metabolites involved in nutrient transport and detoxification pathways in urine samples from women with SGA pregnancies. There were no common SGA-associated metabolite changes identified in plasma and urine.

Our lipidomics study shows evidence of altered glycerophospholipids (GPL) and sphingolipid pathways associated with SGA at 20 weeks of gestation in Cork plasma samples. These two chemical classes of phospholipids are the main lipid class present in biological bilayer, such as cell membranes, and are involved in inflammation, apoptosis, and storage and breakdown of lipids for energy.(Baig et al. 2013) Our findings are in agreement with a metabolomics study performed using samples from the SCOPE study taken at 15 weeks of gestation in Australian participants.(Horgan et al. 2011) In the Horgan *et al.* study, metabolic profiles of maternal plasma, venous cord blood, and plasma samples from a rat model of SGA (reduced uterine perfusion pressure, RUPP, model) were obtained on a UPLC system coupled to a hybrid LTQ-Orbitrap mass spectrometer. Horgan *et al.* identified nineteen metabolites as statistically different in SGA cases as compared to controls; these included phospholipids, such as several lysophosphatidylcholines, phosphatidylcholines, and sphingolipids, most of which were detected in higher levels in SGA groups in maternal plasma samples, thereby supporting our findings of increased plasma glycerophospholipids (GPL) in SGA. In addition, a recent study combining the use of direct infusion LC-MS/MS, proton nuclear magnetic resonance (NMR) and artificial intelligence showed alteration of several pathways in cord blood samples, including phospholipid biosynthesis and fatty acid metabolism, to be associated with SGA (Bahado-Singh et al. 2019).

Our metabolomics study of urine samples taken at 20 weeks of gestation, showed that the metabolites that were significantly altered in SGA are involved in detoxification (D-Glucuronic acid in Auckland samples, Estriol-16-glucuronide in Cork samples), nutrient absorption and transport (Sulfolithocholic acid in Cork samples) pathways. None of the significantly different metabolites of interest were found in

both the Cork and Auckland populations, but several are involved in similar pathways and cellular processes. This may be attributed to the fact that this independent study was run months apart, on a different BEH C18 column therefore chromatographic reproducibility was not minimised. Future projects will take a more targeted approach to metabolomics validation given the inherent problems of chromatographic reproducibility. Nonetheless, differentially expressed metabolites were all detected in lower levels in SGA cases compared to controls in both populations, which suggests detoxification and placental nutrient transport mechanisms maybe less efficient in SGA pregnancies, and warrant further investigation.

Overall, the lipidomics and metabolomics analyses performed on samples taken from Cork women suggests placental insufficiency and inadequate transport of lipids to the placenta, resulting in impaired fetal growth (S. Zhang et al. 2015) detected as early as 20 weeks of gestation. Indeed, the placenta plays a key role in the development of the fetus *in utero*, as it ensures the fetus receives sufficient nutrients – especially oxygen amino acid, glucose and fatty acids (Lager and Powell 2012).

Our study has limitations; one of them is the number of smoking women in Cork population (30 % in the SGA group, and 10 % in the control group), and another is the significant difference of gestational age at delivery between SGA group and controls in Cork population. This latter difference can be attributed to the fact that the participants selected for the SGA group delivered babies with extremely low customised birthweight centile (median of 2.35, IQR 0.8-4.48) compared to the control group (median 49, IQR 35.45-65.83). Cases and controls were matched for age, ethnicity and BMI when the studies were designed, and exclusion of participants would have reduced the power of this study. In addition, using all known risk factors associated with SGA to adjust the data during statistical analysis would have led to overfitting the data. We did, however, use a stringent false discovery rate cut-off at 5 % to select metabolites and lipids of interest, and we further tested the metabolites and lipids of interest to determine if they were significantly correlated with clinical factors (in addition to SGA). No significant correlation was observed with any metabolite of interest and gestational age at delivery, or smoking status, and only

one lipid of interest, CL(72:2), was found to be significantly correlated with smoking status. However, this lipid was also significantly correlated with SGA.

Another limitation of this study is that the analysis aimed to measure changes in lipids concentrations in the maternal circulation. However, it is not fully known whether changes in placenta metabolic pathways affected by SGA would be significantly reflected in the maternal circulation (Duttaroy 2009; H. N. Jones et al. 2007). Yet, an *in vitro* study has shown that human placenta perfused with an 8-fold increase of the free fatty acid docosahexanoic acid (DHA) on the maternal side resulted in a 13-fold increase of DHA transferred to the fetal side (Haggarty et al. 1999). Similarly, for a 2-fold increase of arachidonic acid (AA) on the maternal side, an 8-fold increase in the transfer of AA to the fetal side was measured (Haggarty et al. 1999). These results show that the free fatty acids available in the maternal circulation can have significant impact on the fatty acids transferred to the fetus (Haggarty et al. 1999). In addition, triglycerides are considered the main source of free fatty acids in the placenta and fetus, and are obtained when released from triglycerides by lipoprotein lipase (LPL) (H. N. Jones et al. 2007). Magnusson-Olson and colleagues have shown that chronic exposure (24h) of isolated cytotrophoblasts from term human placentas to increased levels of free fatty acids resulted in decreased LPL activity, but did not alter LPL protein expression (Magnusson-Olsson et al. 2006). These results indicate that impaired placental lipid and fatty acid transport can reflect significant changes in the lipid levels in maternal circulation. However, more experiments are required to determine whether the altered glycerophospholipid levels detected in maternal circulation in the present study could be a direct reflection of placental dysfunction.

In conclusion, this study suggests that higher levels of glycerophospholipids and sphingolipids at 20 weeks of gestation, may be associated with the onset of SGA in participants of the SCOPE study in Cork. The data presented in this study have yielded significant information regarding involvement of glycerophospholipids in the aetiology of SGA. However, whether the correlations represent a cause or an effect of SGA, and the precise mechanistic implications need to be further investigated. Further studies are needed to validate these findings in an independent pregnancy

cohort and to examine whether there may be the potential to use these lipids to predict pregnancies at risk of SGA.

Chapter 4:

Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomics analysis in the Cork pregnancy cohort.

Morillon, A.-C., Yakkundi S., et al. (2020). "Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomic analysis in the cork pregnancy cohort." Metabolomics 16(2): 19. (Morillon et al. 2020b)

4.1. Abstract

Introduction

Preterm birth (PTB) is defined as birth occurring before 37 weeks of gestation, affects 5-9 % of all pregnancies in developed countries, and is the leading cause of perinatal mortality. Spontaneous preterm birth (sPTB) accounts for 31-50 % of all PTB, but the underlying pathophysiology is poorly understood.

Objective

This study aimed to decipher the lipidomics pathways involved in pathophysiology of sPTB.

Methods

Blood samples were taken from SCreening fOr Pregnancy Endpoints (SCOPE), an international study that recruited 5,628 nulliparous women, with a singleton low-risk pregnancy. Our analysis focused on plasma from SCOPE in Cork. Discovery profiling of the samples was undertaken using liquid chromatography-mass spectrometry Lipidomics, and features significantly altered between sPTB (n=16) and Control (n=32) groups were identified using empirical Bayes testing, adjusting for multiple comparisons.

Results

Twenty-six lipids showed lower levels in plasma of sPTB compared to controls (adjusted $p < 0.05$), including 20 glycerophospholipids (12 phosphatidylcholines, 7 phosphatidylethanolamines, 1 phosphatidylinositol) and 6 sphingolipids (2 ceramides and 4 sphingomyelines). In addition, a diaglyceride, DG (34:4), was detected in higher levels in sPTB compared to controls.

Conclusions

We report reduced levels of phospholipids in sPTB. Phospholipid integrity is linked to biological membrane stability and inflammation, while storage and breakdown of lipids have previously been implicated in pregnancy complications. The contribution of phospholipids to sPTB as a cause or effect is still unclear; however, our results of

differential plasma phospholipid expression represent another step in advancing our understanding of the aetiology of sPTB. Further work is needed to validate these findings in independent pregnancy cohorts.

4.2. Introduction

Preterm birth (PTB) is defined as birth occurring before 37 weeks of gestation and can be spontaneous or iatrogenic. Preterm delivery rates vary between countries, but are estimated to be 5-9 % of pregnancies in Europe and other developed countries, with rates still rising in most industrialised countries (Slattery and Morrison 2002; Robert L. Goldenberg et al. 2008). PTB accounts for 75 % of perinatal deaths, with two thirds of these deaths occurring in preterm infants delivered before 32 weeks of gestation (Slattery and Morrison 2002). Spontaneous preterm birth (sPTB) is estimated to occur in 31-50 % of all PTB (Slattery and Morrison 2002) and the pathways involved are not yet fully understood. A previous study of low-risk nulliparous pregnant women in an international cohort, SCreening fOr Pregnancy Endpoints (SCOPE) study (www.scopestudy.net), used clinical data to determine risk factors linked to two sub-phenotypes of sPTB, sPTB with intact membranes (sPTB-IM) and sPTB occurring after prelabour rupture of membranes (sPTB-PPROM). Characteristics such as marijuana use before conception, shorter cervical length, or strong family history of low birth weight, were linked to higher risks of sPTB. However, the models built using the reported risk factors were not robust enough to predict sPTB (Dekker et al. 2012). Indeed, the area under the curve (AUC) of the model for sPTB-IM was 0.69, with a 0.39 sensitivity based on 90 % specificity; the model for sPTB-PPROM showed an AUC of 0.79, with 0.49 sensitivity based on 90 % specificity. As a result, research efforts have focused on finding blood based biomarkers for early diagnosis and intervention.

Metabolites are intermediate compounds involved in biochemical processes and are a reflection of cellular metabolic pathways of the organism. Similarly, lipids are a subset of metabolites which form the matrix of our cell membranes and support a variety of biological functions such as energy storage, cell metabolism, and cell

signalling. Together, metabolites and lipids levels are regulated in response to environmental modifications (Goodacre et al. 2004), and as such, are a direct reflection of the phenotype (Patti et al. 2012); this makes them ideal candidates for biomarkers and therapeutic targets. Several recent studies have used metabolomics methods for biomarker discovery of PTB in plasma (Lizewska et al. 2018), in amniotic fluid samples (Romero et al. 2010; Menon et al. 2014), and in cervicovaginal secretions (Auray-Blais et al. 2011). A systematic review performed on the literature between 1965 and 2008 showed that various sample types have been used for the study of biomarkers of spontaneous preterm birth (Menon et al. 2011), and 46.2 % of included studies used maternal blood samples, 31.0 % used amniotic fluid samples, and 1.2 % used urine samples. In addition, the collection of blood sample is minimally invasive compared to the collection of amniotic fluid for example, which can present risks for the fetus (Beta et al. 2018).

In recent years, several studies have linked altered levels of lipids and higher risk of adverse pregnancy outcomes (Baig et al. 2013; Jiang et al. 2017; Reece et al. 1997). Baig *et al.* have demonstrated that altered levels of glycerophospholipids and sphingolipids in syncytiotrophoblast microvesicles shed from the placenta, are associated with increased risk of preeclampsia or recurrent miscarriages (Baig et al. 2013). Reece *et al.* have shown an association between preterm birth (<37 weeks of gestation) and higher levels of arachidonic acid and docosapentaneoic acid levels in maternal blood samples taken at the time of delivery (Reece et al. 1997). Jiang et al. preformed a systematic review aiming to analyse the associations between altered maternal lipids levels (cholesterol, triglycerides, high and low lipoprotein-cholesterol), and the risk of preterm birth (Jiang et al. 2017). Their meta-analysis showed that altered maternal levels of lipids are associated with a higher risk of preterm birth, for either higher levels (pooled OR for elevated levels 1.68[95 % CI 1.24-2.26]), or lower levels (pooled OR 1.52[95 % CI 1.13-3.82]).

To date, no study has identified metabolites or lipids robust enough to be used as biomarkers of PTB, as they need to be validated in different populations, using larger sample sizes. Here we describe the first discovery lipidomics study to be undertaken in sPTB. The aim of this discovery profiling study was to help determine the clinically

relevant lipidomic pathways involved in the aetiology of sPTB. For this purpose, lipidomics profiles were obtained from plasma samples taken at 20 weeks of gestation from women recruited to the SCOPE study in Cork, Ireland (Louise C Kenny et al. 2014).

4.3. Materials and Methods

4.3.1. Study participants

Flow chart of participant selection and samples preparation are presented in **Figure 4.1**. Selected participants were women recruited through the SCOPE study (www.scopestudy.net), in Cork, Ireland. SCOPE is an international longitudinal birth cohort that recruited over 5,000 healthy nulliparous women with singleton pregnancies in Australia, New Zealand, the United Kingdom and Ireland (Louise C Kenny et al. 2014). Written informed consent and ethical approval were obtained prior to sample collection (ECM5 (10) 05/02/08). In Cork SCOPE sub-cohort, 1773 women were recruited, including 56 women with spontaneous preterm birth before 37 weeks of gestation, which represents approximately 3 % of this population. In the present cross-sectional study from a prospectively collected birth cohort, we chose to focus on women who had an "early" spontaneous preterm birth, defined as spontaneous preterm birth <34 weeks of gestation (n=16, 0.9 % of Cork SCOPE population). Plasma samples were collected at 20 weeks of gestation, and women who later delivered prematurely before 34 weeks of gestation (cases) were matched to women who had an uncomplicated pregnancy (controls). Controls (n=32) were matched to cases (n=16) according to age (+/- 5 years), and body mass index (BMI, +/- 3.0 kg/m^2). Detailed clinical and demographic data is presented in **Table 4.1**.

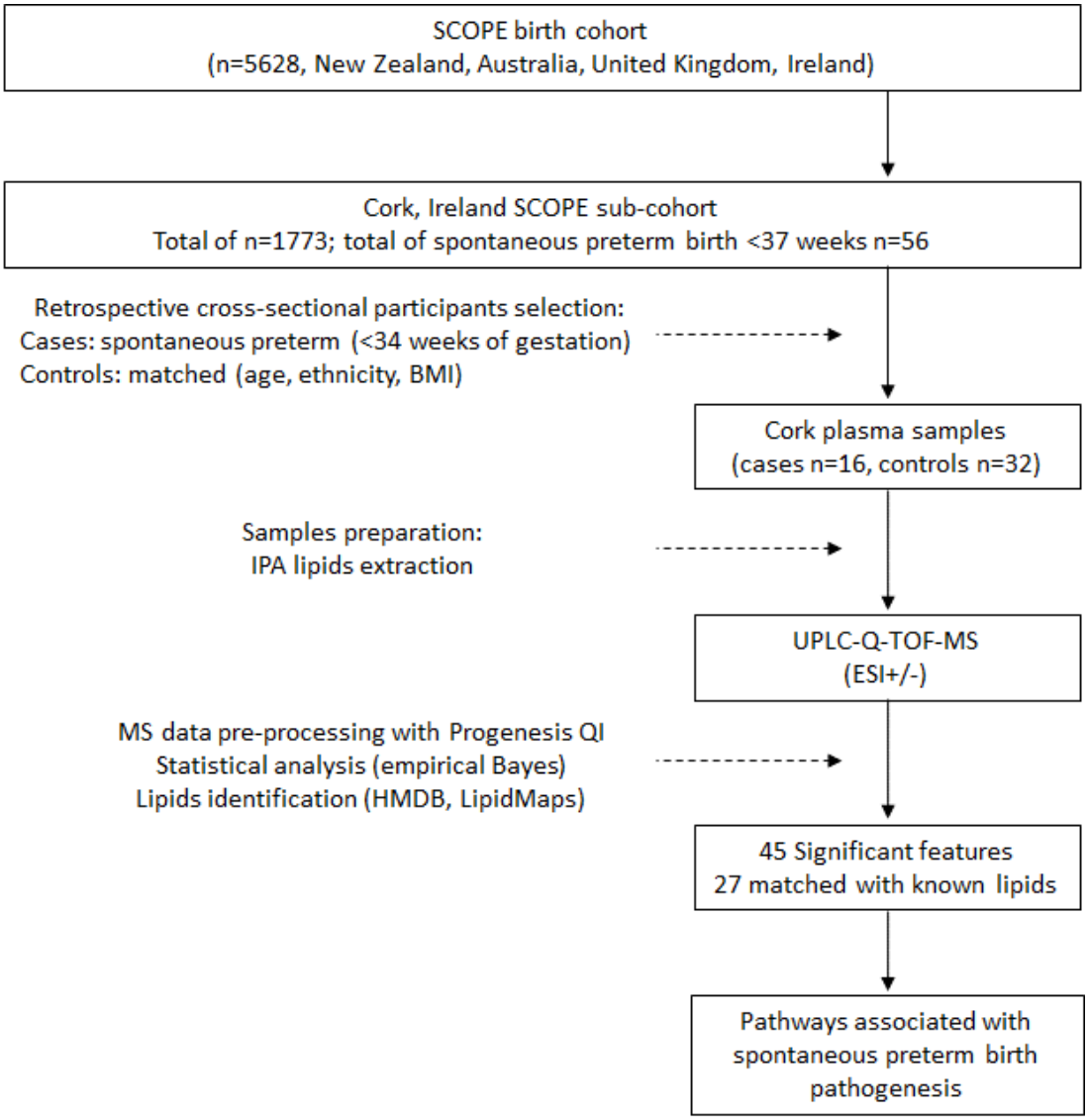


Figure 4.1: Flow chart of participant selection, samples preparation and analysis.

Table 4.1: Demographic and clinical data of the study population.

	Control (n=32)	Case (n=16)	p-value
Age (years)	29.69 (5.03)	30.31 (6.93)	0.723
BMI, grouped according to World Health Organisation			
0 - <18.5, underweight	0	1 (6.2)	0.555
≥18.5 and <25, normal weight	16 (50.0)	7 (43.8)	
≥25 and <30, overweight	8 (25.0)	4 (25.0)	
≥30, obese	8 (25.0)	4 (25.0)	
Ethnicity			
Caucasian	32 (100)	16 (100)	NA
Marital status			
Single	1 (3.1)	6 (37.5)	0.006
Married	21 (65.6)	7 (43.8)	
Stable relationship	10 (31.3)	3 (18.8)	
Years of schooling	14 (13-14)	13.5 (13-14)	0.801
Years of schooling, in groups			
<12 years	0	2 (12.5)	0.120
12 or 13 years	15 (46.9)	6 (37.5)	
>13 years	17 (53.1)	8 (50.0)	
Job situation			
Full time work	25 (78.1)	12 (75.0)	0.130
Part time work	4 (12.5)	0	
Homemaker	2 (6.3)	0	
Unemployed	1 (3.1)	2 (12.5)	
Student	0	1 (6.3)	
Sickness beneficiary	0	1 (6.3)	

Socioeconomic index	45.66 (19.02)	35.81 (13.31)	0.044
Smoking status at 1st visit (15 weeks)			
Never smoked	23 (71.9)	10 (62.5)	
Smoked pre-pregnancy, but quit smoking before pregnant	0	0	
Smoked in pregnancy, but quit before 1st visit	6 (18.8)	3 (18.8)	0.642
Smoking at 1st visit (15 weeks)	3 (9.4)	3 (18.8)	
Alcohol status at 1st visit (15 weeks)			
Never consumed alcohol	3 (9.4)	3 (18.8)	
Consumed alcohol pre-pregnancy, but quit before conception	4 (12.5)	1 (6.3)	
Consumed alcohol in pregnancy, but quit before 1st visit	20 (62.5)	9 (56.3)	0.734
Continuing to drink at 1st visit	5 (15.6)	3 (18.8)	
Birthweight (g)	3560 (3385-3775)	1665 (1102.5-2070)	< 0.001
Customised birthweight centile [a]	47.6 (34.18-81.1)	15.1 (9.03-73.85)	0.076
Gestational age at delivery (weeks)	40.50 (39.79-41.40)	32.29 (28.75-33.36)	< 0.001

Selected participants were women recruited as part of SCreening fOr Pregnancy Endpoints study (SCOPE www.scopestudy.net), in Cork, Ireland. Cases were defined as women who delivered spontaneously before 34 weeks of gestation, and were matched according to body mass index, age and ethnicity to women who delivered at term (controls). Values are shown as mean (SD), median (interquartile range) or n (%). [a] customised birthweight centile: adjusted for mother's height, weight at 15 weeks visit, ethnicity, sex and weight of baby and gestation at delivery of baby; BMI: Body mass index; SGA: small for gestational age; NA: not applicable; all mothers were nulliparous.

4.3.2. Reagents and materials

Liquid chromatography grade iso-propanol (IPA), acetonitrile (ACN) and ammonium formate were purchased from Fisher Scientific (Loughborough, UK). LC-MS glass vials and ultra-performance liquid chromatography (UPLC) columns were purchased from Waters (Waters, Wexford, Ireland).

4.3.3. Sample preparation

Heparinised plasma samples taken from participants at 20 weeks of gestation were randomised before extraction. Samples were taken out of -80°C storage and allowed to thaw on ice, before being transferred to labelled Eppendorf tubes (200 µl). Lipids were extracted as previously described (Sarafian et al. 2014), iso-propanol (IPA) chilled at -20°C was added (600 µl) to the plasma. Samples were then vortex mixed for 1min and incubated for 10 min at room temperature, then stored at -20°C overnight to improve protein precipitation. The following day, samples were centrifuged at 14,000 g for 20 min at room temperature. For each sample, the supernatant was transferred in correctly labelled LC vials. For quality control, a volume of 30 µl was taken from each sample, pooled in a tube and vortexed to create pooled quality control samples (QC). A volume of 100µl of pooled samples were aliquoted in different QC vials.

4.3.4. Global Lipidomics profiling analysis

Samples were analysed using an ultra-high performance liquid chromatography quadrupole time-of-flight (UPLC-Q-TOF) mass spectrometry. The UPLC system was a Waters ACQUITY system (Waters Corp, Wilmslow, UK), coupled with a BEH C18, 1.7 µm, 2.1 x 100 mm analytical column (Waters Corp, Wexford, Ireland). The samples were analysed in a randomised order, and as technical triplicates, with an injection volume of 4 µl (ESI+) and 7 µl (ESI-). Pooled quality control samples (QC) were injected to condition the column (n = 8 injections) prior to the start of the analysis, and thereafter every tenth injection. A 23 min gradient elution was applied at a flow rate of 0.4 ml/min using two mobile phases, a mix of ACN and water (60:40, v:v) with 10

mM of ammonium formate (A), and a mix of IPA and ACN (90:10, v:v) with 10 mM of ammonium formate (B). The elution gradient was as follows: initial conditions at 30 % B; from 1 to 15 min increased up to 99 % B; from 15 to 20 min, maintained at 99 % B; from 20 to 22 min decreased to 30 % B; from 22 to 23 min, returned to initial conditions of 30 % B. During the analysis, samples were maintained at 4 °C and the column at 65 °C Mass spectrometry analysis was performed using a Synapt G2-S Q-ToF (Waters Corp, Wilmslow, UK) with data collected in continuum format using positive and negative electrospray ionisation (ESI +/-). The data independent acquisition (DIA) mode, MS^e (Bateman et al. 2002; Silva et al. 2005) was used for data acquisition. Data were acquired from 50 to 1,500 m/z range, in resolution mode. Precursor (low energy) and fragment (high energy) ion data were collected within the same acquisition with a scan time of 0.1 s for each, providing a total cycle time of 0.2 s. In the case of high energy, a linear collision energy ramp (20-40 eV) was applied over the 0.1 s scan. Capillary voltage was set to 1.5 kV, sampling cone to 30 V and extraction cone to 5 V. The source was set at 120°C, and desolvation temperature at 650 °C. Desolvation gas flow rate was set at 800 L/h and cone gas at 50 L/h. Detector set up was performed using leucine enkephalin (LeuEnk) (Waters, Wexford, Ireland) and SYNAPT was calibrated using sodium formate (Waters, Wexford, Ireland), according to manufacturer's instructions. Real time lock-mass correction was achieved by infusing LeuEnk at 10 µL/min through a lock-spray probe and acquired every 30 s.

4.3.5. Data pre-processing

MS^e data were processed using Progenesis QI version 2.4 (Nonlinear dynamics, Newcastle, UK). Progenesis QI is a data analysis software package that accurately processes LC-MS data, based on accurate mass measurements, isotopic fit and fragmentation data. Data were chromatographically aligned using an appropriate pooled QC as the alignment reference. Data were peak picked and normalised to all compounds. Adducts corresponding with M+H, $M+H-H_2O$, $M+H-2H_2O$, $M+NH_4$, M+Na, M+K, M+2H, M+2Na were considered during peak picking. Compound

measurements (raw and normalised compound abundances) were exported from Progenesis QI for downstream statistical analysis.

4.3.6. Statistical analysis

For statistical analysis of the demographic and clinical data we applied the Student T test, Mann-Whitney U test, or Pearson χ^2 test, and multiple comparisons corrections as appropriate (IBM SPSS Statistics 24). Results were considered statistically significant if the p-value was less than 0.05 (**Table 4.1**). Randomisation of samples before preparation was performed using block randomisation based on the patients' BMI and the outcome. No significant dependency between measurement order, the outcome and biometric and clinical information about patients was observed ($p<0.05$; Mann Whitney U test, Spearman correlation, Chi square test and Kruskal-Wallis test as applicable; Benjamini and Hochberg step-up FDR-controlling procedure) (Benjamini and Hochberg 1995).

Statistical analyses of UPLC-MS data were performed using the R statistical software (R Core Team 2013) and the Bioconductor package limma (Ritchie et al. 2015). UPLC-MS analyses were performed with two electrospray ionisation (ESI) modes (ESI+ and ESI-), which differ significantly in their analytical sensitivity and technical variability and therefore analysed independently. Briefly, median normalisation was performed on the raw UPLC-MS data to correct for the systematic measurement biases known to be associated with the UPLC-MS platform used. Quality control procedures were applied to check for systematic measurement biases (Broadhurst et al. 2018). Measurement precision was checked for each feature by computing the missing rate and coefficient of variation (CV) over the replicate measurements. Features with a missing rate greater or equal to 20 % and features with a CV greater or equal to 30 % were not considered in subsequent analyses. Features predictive for the outcome were selected using the empirical Bayes method, adjusting the design for replicate measures (Smyth 2004). Empirical Bayes methods have been successfully applied on LC-MS data (Kammers et al. 2015; Margolin et al. 2009), and offer the advantage of correcting for false discovery rate while identifying differential features. This approach was verified by comparing them to feature selection made using Mann Whitney U test, on the average measurement per patient with multiple testing

correction. Agreement between the two methods was demonstrated (Data not shown).

4.3.7. Putative annotation of lipids

The exact mass of features with a significant adjusted value ($p < 0.05$) were searched against the Human Metabolome Database version 4.0 (David S. Wishart et al. 2018), Lipid Maps version of January 2019 (Cotter et al. 2007) and Chemspider (searching the databases ChEBI version of February 2018, and KEGG version of April 2014; http://www.chemspider.com), using the Progenesis QI identification tool. Search parameters were set for an exact mass tolerance of 5 and 10 ppm, for precursor and fragment ions respectively. Putative annotations were then manually verified when fragmentation data was available, as shown in **Appendix C**. In accordance with the MSI reporting standards, we have achieved metabolite identification level 2, or putatively annotated compounds (Salek et al. 2013).

4.4. Results

We undertook lipidomics analysis of plasma samples from women participating in the SCOPE pregnancy cohort in Cork, Ireland. Samples were collected at 20 weeks of gestation, and selected cases (delivery before 34 weeks of gestation) were matched to controls according to age, BMI and ethnicity (**Table 4.1**). The clinical data analysis showed that women who had a spontaneous preterm birth were more likely to be married (43.5 %), or single (37.5 %), while only 18.8 % were in a stable relationship. The ESI + data set generated for analysis included 5,878 features, and principal component analysis (PCA) and volcano plots are shown in **Figure 4.2a** and **4.2c**. After data pre-processing, data filtration (as described above), statistical analysis was performed using empirical Bayes methods on 4,845 features. We putatively annotated 45 features as significantly differentially expressed between cases and controls, of which 27 remained significant following correction for multiple testing (**Table 4.2**). Of these 27 annotated lipids, 26 are phospholipids: 20 glycerophospholipids (12 phosphatidylcholines (PC), 7 phosphatidylethanolamines

(PE), and 1 phosphatidylinositol (PI)), and 6 sphingolipids (2 ceramides and 4 sphingomyelins (SM)). Notably, all phospholipids were significantly lower in cases compared to controls. The last metabolite identified was a diglyceride, (DG (34:4)), detected at higher levels in sPTB compared to controls. Box plots of significant metabolites annotated in the Cork cohort are represented in **Appendix C**. The ESI - data set generated a total of 769 features, PCA and volcano plots are shown in **Figure 4.2b and 4.2d**, and after data pre-processing, data filtration, and statistical analysis, no features were found to be significantly altered in sPTB group compared to control group.

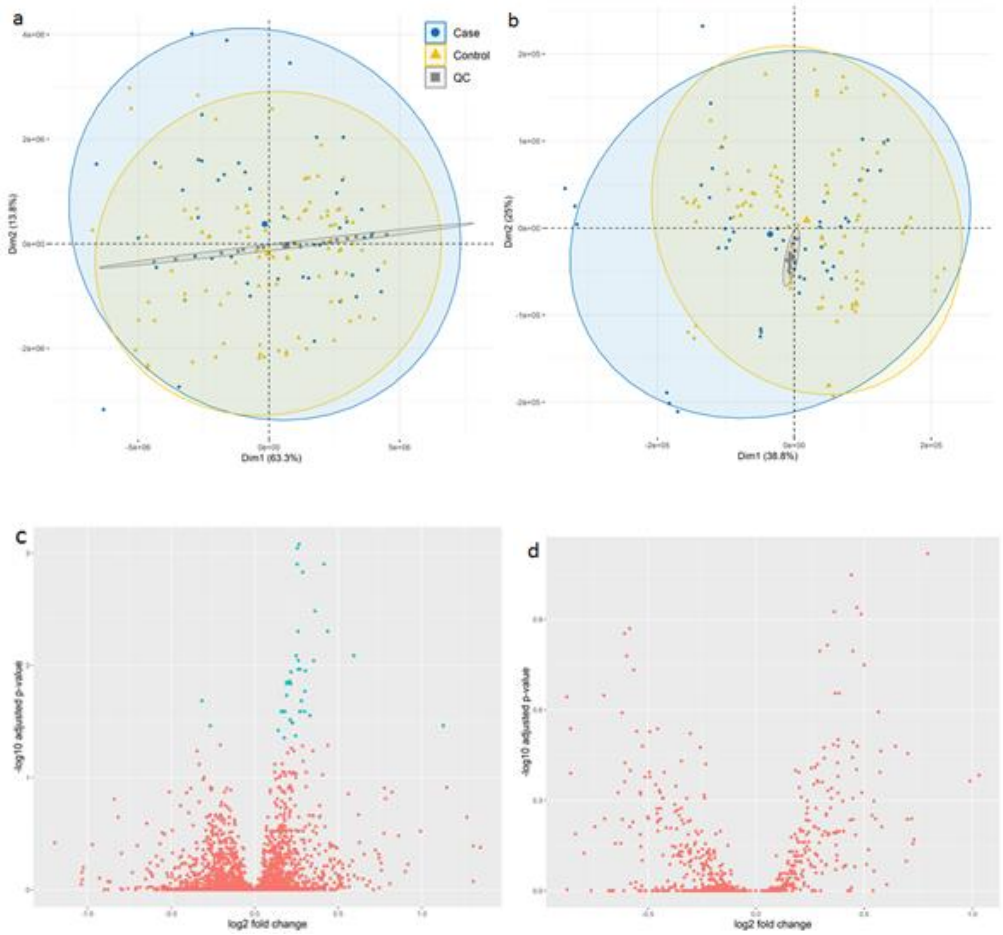


Figure 4.2: PCA score plot of all samples in ESI+ (a) and in ESI- (b) and respective volcano plots (c, and d). In PCA plots, blue dots represent sPTB case samples, yellow triangles represent control samples, and grey squares represent QC samples. Volcano plots show $\log_2$ fold change plotted against $-\log_{10}$ of adjusted p-values in ESI+ (c) and ESI- (d), plotted in blue are the variables with an adjusted p-value <0.05.

Table 4.2: Putatively annotated lipids that were significantly (p<0.05) altered in spontaneous preterm (sPTB) group compared to control.

Identification	Chemical class	p-value	Adjusted p-value	Direction	Fold Change (95% CI)
PC(35:1)	Glycerophospholipid	2.06E-05	**1.46E-02**	DOWN	1.23 (1.11-1.39)
PC(40:5)	Glycerophospholipid	1.30E-04	**2.43E-02**	DOWN	1.26 (1.10-1.45)
PC(42:4)	Glycerophospholipid	5.53E-08	**7.11E-04**	DOWN	1.19 (1.12-1.26)
PC(42:5)	Glycerophospholipid	2.75E-05	**1.21E-02**	DOWN	1.15 (1.07-1.27)
PC(O-38:4)	Glycerophospholipid	5.40E-06	**1.22E-02**	DOWN	1.16 (1.07-1.26)
PC(O-42:4)	Glycerophospholipid	7.51E-06	**7.51E-03**	DOWN	1.20 (1.09-1.26)
PC(O-44:6)	Glycerophospholipid	4.68E-05	**2.14E-02**	DOWN	1.13 (1.06-1.19)
PC(P-36:1)	Glycerophospholipid	8.61E-05	**2.73E-02**	DOWN	1.17 (1.07-1.26)
PC(P-38:3)	Glycerophospholipid	6.33E-05	**2.14E-02**	DOWN	1.13 (1.07-1.20)
PC(P-40:3)	Glycerophospholipid	8.19E-05	**2.14E-02**	DOWN	1.23 (1.14-1.36)
PC(P-40:4)	Glycerophospholipid	2.75E-05	**2.11E-02**	DOWN	1.12 (1.06-1.18)
PC(P-40:5)	Glycerophospholipid	2.79E-08	**6.57E-04**	DOWN	1.20 (1.12-1.29)
PE(38:0)	Glycerophospholipid	9.60E-06	**9.12E-03**	DOWN	1.21 (1.11-1.29)
PE(38:1)	Glycerophospholipid	3.26E-05	**3.69E-02**	DOWN	1.13 (1.06-1.21)
PE(38:2)	Glycerophospholipid	7.61E-05	**2.14E-02**	DOWN	1.21 (1.09-1.34)
PE(40:3)	Glycerophospholipid	1.31E-05	**9.56E-03**	DOWN	1.23 (1.12-1.35)
PE(40:4)	Glycerophospholipid	1.70E-05	**9.68E-03**	DOWN	1.16 (1.09-1.26)
PE(42:6)	Glycerophospholipid	9.97E-06	**7.12E-03**	DOWN	1.51 (1.27-1.78)
PE-Cer(38:3)	Glycerophospholipid	3.34E-06	**2.71E-02**	DOWN	1.17 (1.11-1.22)
PI(35:0)	Glycerophospholipid	4.83E-05	**2.99E-02**	DOWN	2.19 (1.53-3.17)
CerP(44:1)	Sphingolipid	7.17E-05	**2.14E-02**	DOWN	1.21 (1.10-1.30)
GalactosylCer(42:2)	Sphingolipid	2.48E-07	**1.07E-03**	DOWN	1.19 (1.12-1.27)
SM(32:2)	Sphingolipid	7.37E-05	**2.58E-02**	DOWN	1.16 (1.08-1.26)
SM(34:1)	Sphingolipid	1.25E-04	**4.63E-02**	DOWN	1.17 (1.09-1.25)
SM(35:1)	Sphingolipid	9.00E-07	**7.68E-03**	DOWN	1.28 (1.17-1.40)
SM(38:2)	Sphingolipid	9.68E-06	**1.21E-02**	DOWN	1.14 (1.08-1.20)
DG(34:4)	Fatty acyl	1.29E-04	**2.99E-02**	UP	1.20 (1.08-1.33)

Cer: Ceramide; CerP: Ceramide-1-phosphate; DG: diglyceride; PC: Phosphatidylcholine; PE: Phosphatidylethanolamine; PE-Cer: Ceramide Phosphoethanolamine; PI: Phosphatidylinositol; SM: Sphingomyeline. **Bold**: adjusted p-value <0.05. Fold change reported with 95 % confidence interval.

4.5. Discussion

Our study provides the first evidence that altered plasma phospholipids expression at 20 weeks of gestation is associated with the onset of sPTB, and our findings are of broad relevance to "at risk" pregnancy states. Sphingolipid metabolism and glycerophospholipid metabolism are involved in pertinent molecular processes such as storage and breakdown of lipid molecules for energy, apoptosis, inflammation, and cell-membrane stabilisation (Baig et al. 2013). The sphingolipid and glycerophospholipid pathways are also interconnected and of mutual dependence, therefore it is plausible that an overall reduction in phospholipids are implicated in the pathophysiology of sPTB (**Figure 4.3**). Alterations in phospholipid metabolism have previously been associated with pregnancy complications (Horgan et al. 2011; Baig et al. 2013) and animal models of pregnancy loss (Mizugishi et al. 2007), and our study represents a further step in understanding the involvement of phospholipid metabolism in sPTB. Future measures of altered phospholipid metabolism such as ELISA based technology, and integrated multi-omics techniques should be investigated – all of which hold promise for developing prognostic biomarkers for at risk pregnancy states.

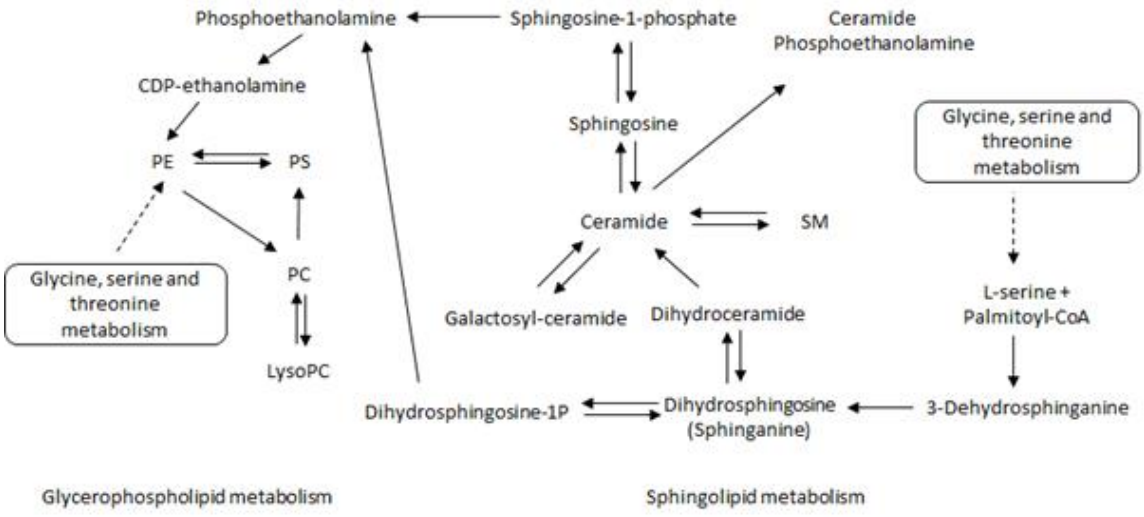


Figure 4.3: Simplified overview of Glycerophospholipid and Sphingolipid metabolisms PC: Phosphatidylcholine; PE: Phosphatidylethanolamine; PS: Phosphatidylserine; SM: Sphingomyelin. Palmitoyl-CoA: Palmitoyl Coenzyme A.

Using a unique prospective cohort, we investigated blood plasma samples at 20 weeks of gestation, from women whose pregnancies ended in sPTB (< 34 weeks of gestation) in comparison to matched controls, i.e. women who had uncomplicated pregnancies. We specifically putatively annotated 20 glycerophospholipids and 6 sphingolipids as differentially expressed following correction for multiple testing, and the uniform decreased expression of these two phospholipid groups was particularly striking in the sPTB group (**Figure 4.2**). In support of our observations in sPTB in SCOPE Cork, a recent metabolomic profiling study in the same cohort of women implicated bile acids, prostaglandins and Vitamin D as differentially expressed at 15 and 20 weeks of gestation (E. C. Considine et al. 2019). This independent study performed at the University of Manchester, profiled metabolites using alternative extraction protocols (Dunn et al. 2011a) and LC-MS/MS analyses. Despite different preparation methods and analytical analyses, both studies report increased fatty acids levels and reduced diacylglycerophosphoinositol in sPTB. Together, these studies provide complementary evidence for the implication of metabolomic and lipidomic pathways in the pathophysiology of sPTB and suggest dysregulation on a systems level.

In further support of altered lipid profiles in adverse pregnancy outcomes, Baig et al. (2013) reported altered glycerophospholipids and sphingolipids in placental syncytiotrophoblast microvesicles, in preeclampsia, which are shed from the placenta into maternal circulation (Baig et al. 2013). Evidence from animal models of at risk pregnancy also implicate altered lipid metabolism, whereby Mizugishi and colleagues (2007) reported that a critical enzyme of sphingolipid metabolism (sphingosine kinase, Sphk) was knocked out, resulting in early pregnancy loss in mice (Mizugishi et al. 2007). Finally, our findings of altered phospholipids in sPTB may be broadly related to the pathophysiology of obstetric antiphospholipid syndrome (OAPS), an auto-immune disease consistently associated with pregnancy complications. OAPS is characterised by presence of circulating antiphospholipid antibodies (aPL) (Antovic et al. 2018). These aPLs target a circulating blood protein, β_2-glycoprotein I (β_2-GPI), that has high affinity and binds to phospholipids present on the surface of cell membranes, and in doing so create a cascade of events that can

lead to thrombosis, vasculopathy, inflammation, and a range pregnancy complications (Garcia and Erkan 2018). A recent systematic review aiming to examine all biomarkers of sPTB in human studies, from 1965 to 2008 (Menon et al., 2011), showed that 55 % of the biomarkers studied are involved with inflammation and immune response, and that interleukin 6 (IL6), IL8 and corticotrophin-releasing hormone (CRH) were the most investigated biomarkers. The other biomarkers studied were involved in several other pathways, such as stress, hormonal metabolism, cellular metabolism, apoptosis, and placental dysfunction. In addition, proteins are the most studied type of biomarker in the articles included in this systematic review. The authors noted that none of these biomarkers showed the ability to predict sPTB accurately or help in better understanding the pathophysiology of this pregnancy adverse outcome. This shows how metabolomics based approaches could be better suited as biomarkers, since metabolites are a direct reflection of the cellular activity (Patti et al., 2012).

Our study has limitations, in particular the low number of sPTB samples with a delivery before 34 weeks of gestation, which is attributed to the fact that the participants recruited for SCOPE were low-risk nulliparous women. Nevertheless, our study was well designed with a 1:2 case:control ratio, participants were carefully matched for BMI, age and ethnicity, and we applied multiple testing to identify 27 significant features following multiple testing correction. Moreover, no additional findings were determined from the remaining metadata. Overall, our findings warrant replication in larger independent cohorts, using multi-omics measures to confirm dysregulation on a systems level, and explore whether the contribution of these 27 altered lipids is a cause or effect of preterm birth.

In conclusion, this study reports reduced glycerophospholipids and sphingolipids at 20 weeks of gestation, prior to the onset of sPTB in women in the Cork SCOPE study. Our findings confirm and extend findings by Considine et al. (2019), suggesting convergent metabolomic and lipidomic pathways in the same cohort of women (E. C. Considine et al. 2019), which may serve as potential therapeutic targets or putative markers for the prediction of sPTB. Further research is needed to validate these findings in independent pregnancy cohorts.

Chapter 5:

Altered Plasma Levels of Phosphatidylcholine and Triacylglycerol in Preterm Birth; A Biomarker Discovery and Validation study in the SCOPE Birth Cohort.

Kirsten G Dowling; Aude-Claire Morillon *et al.* Manuscript prepared for *American Journal of Obstetrics and Gynecology*.

5.1. Abstract

Introduction

The pathophysiology of spontaneous Preterm Birth (sPTB) remains unclear, and its accurate prediction in clinical practice is important as early intervention is associated with improved outcome. We undertook a discovery lipidomic study at 15 weeks of gestation, and validation in longitudinal plasma samples taken at 15 and 20 weeks of gestation.

Methods and Findings

Participants were selected from the international SCreening fOr Pregnancy Endpoints (SCOPE) study, which recruited low-risk nulliparous women. The discovery analysis focused on participants from Cork (Ireland) and Auckland (New Zealand) at 15 weeks of gestation using samples from women whose pregnancies ended in sPTB (cases, n=39), matched to control women who had an uncomplicated pregnancy resulting in term delivery (controls, n=39). Using liquid chromatography-quadrupole time of flight mass spectrometry (LC-QTOF), we identified 122 features differentially expressed between sPTB cases and controls using empirical Bayes method (FDR <0.05), with 77 phospholipids found to be significantly reduced in the sPTB group. Five lipids, including a phosphatidylcholine (PC), phosphatidylethanolamine (PE), diacylglycerol (DG), triacylglycerol (TG) and a ceramide (Cer), were selected for targeted validation in plasma samples from Cork participants at both 15 weeks of gestation (cases n=54, controls n=108) and 20 weeks of gestation (cases n=54, controls n=108), using liquid chromatography-triple quadrupole mass spectrometry. To assess longitudinal trends throughout the course of pregnancy, we also examined the expression of these lipids in a small number of plasma samples taken at time of delivery.

We confirmed two lipid targets as differentially expressed between sPTB and controls: PC (15:0/22:6) showed reduced levels in sPTB at 15 weeks of gestation (p=0.015), and TG (18:3/18:2/18:3) showed increase levels in sPTB at 20 weeks (p=0.002). Two lipids (PE (P-16:0/20:5), DG (14:0/18:1) showed no significant

differences at either 15 or 20 weeks of gestation, and Cer (m18:0/22:0) was eliminated as a biomarker due to poor reproducibility (CV >30 %) .

Conclusions

Targeted validation of putative biomarkers is a major bottle neck in the biomarker discovery field, particularly due to the lack of suitable cohorts of samples for biomarker validation. Our biomarker discovery analysis at 15 weeks of gestation in Cork and Auckland cohorts showed a trend towards decreased glycerophospholipids expression in sPTB cases in comparison to controls. Our biomarker validation analysis confirmed altered PC (15:0/22:6) and TG (18:3/18:2/18:3) at 15 and 20 weeks in sPTB, respectively. Although the targeted PC and TG lipids were not strong individual predictors of sPTB, our findings broadly implicate phospholipid metabolism in the pathophysiology of sPTB. Furthermore, our study highlights the challenges in selecting putative biomarkers to take forward for validation, and the importance of longitudinal molecular profiling when screening for pregnancy endpoints such as sPTB.

5.2. Introduction

An estimated 75 % of perinatal mortality and more than half of long-term neonatal morbidity are attributed to preterm birth (PTB) (McCormick 1985) PTB is defined as a birth occurring before 37 weeks of gestation and can be spontaneous or iatrogenic. PTB is estimated to occur in 5-9 % of pregnancies in Europe, and trends show that the rates are rising in most industrialised countries (Robert L. Goldenberg et al. 2008; Slattery and Morrison 2002). Spontaneous preterm birth (sPTB) is estimated to be responsible for 31 to 50 % of all cases of PTB (Slattery and Morrison 2002). Despite increasing research interest, the pathways involved in its aetiology are still not fully understood, and there is no reliable clinical test capable of predicting sPTB.

Metabolites are small molecules (i.e. amino acids, nucleotides, or lipids), used as intermediate compounds, and involved in all biological processes. Lipids are involved in a wide variety or pathways, including pathophysiological processes, by being used as signalling molecules, as elements of cell membranes, and as a source of energy

(H.-C. Lee and Yokomizo 2018). In response to environmental changes, the levels of metabolites, and of its lipid subset, are modified (Goodacre et al. 2004), which makes these molecules a direct reflection of the phenotype (H.-C. Lee and Yokomizo 2018; Patti et al. 2012), and ideal candidates for biomarkers in the prediction of disease.

Several studies performing metabolomics for biomarker discovery of sPTB using plasma or amniotic fluid samples (Morillon et al. 2020b; Romero et al. 2010; Menon et al. 2014) have been published in recent years, but to our knowledge no validation of potential biomarkers has been undertaken. In general, problems with lack of suitable cohorts for validation, poor reproducibility due to the nature of HPLC analysis, matrix effect, accurate metabolite/lipid feature identification, have hindered biomarker validation studies (Dunn et al. 2011b; Khamis et al. 2020). The aim of the present study was to undertake biomarker discovery and validation to identify plasma lipid biomarkers of sPTB, from pregnant women in the SCOPE cohort.

5.3. Materials and Methods

Analytical standards SPLASH® Lipidomix® from Avanti® Polar Lipids (Alabaster, Alabama, USA) and N-C24:1-deoxysphinganine, 1 mg (Cer (m18:0/24:1)) were purchased from Merck (Arklow, Ireland). Liquid chromatography grade water, isopropanol (IPA), acetonitrile (ACN) and ammonium formate were purchased from Fisher Scientific (Loughborough, UK). Autosampler glass vials and ultra-performance liquid chromatography (UPLC™) columns, leucine encephalin and sodium formate were purchased from Waters (Wexford, Ireland). Praxair provided argon (Praxair Gases Ireland Ltd., Immingham, UK), and nitrogen was generated using a nitrogen generator (Peak Scientific Instruments Ltd., Inchinnan, UK.)

5.3.1. Study participants

Selected participants were women recruited into the SCreening fOr Pregnancy Endpoints (SCOPE) study, an international pregnancy cohort (www.scopestudy.net). Over 5,000 low-risk, nulliparous women, with singleton pregnancies were recruited

in Australia, Ireland, New Zealand and United Kingdom [12]. Our study focused on participants from Cork (Ireland), and Auckland (New Zealand). We chose these two cohorts because they were similar in terms of age and socioeconomic profiles but geographically and temporarily distinct. Written informed consent and ethical approvals were obtained prior to sample collection in Ireland (ECM5 (10) 05/02/08) and New Zealand (AKX/02/00/364). In the Cork population, 1773 women were recruited to the SCOPE study, 56 (3 %) of which had a spontaneous preterm birth before 37 weeks of gestation. In the Auckland population, a total of 2034 women were recruited, including 85 (4 %) women who had a spontaneous preterm birth before 37 weeks of gestation. Cases were matched according to age (±5 years) and body mass index (BMI, ±3.0 kg/m^2) to control women who had no preterm birth, preeclampsia, or SGA complications (**Table 5.1**).

For the initial discovery lipidomic screen, a subset of these sPTB cases where preterm birth occurred at <34 weeks of gestation was selected for analysis. This subset consisted of heparinised plasma collected at 15 weeks of gestation of 16 cases:16 controls from Cork and 23 cases:23 controls from Auckland (n=78 total samples). The targeted validation of lipid markers was performed on plasma samples collected at 15 and 20 weeks of gestation from 54 cases and 108 controls from the Cork SCOPE cohort. Cases and controls were matched by BMI and age in a 1:2 case:control study design. The 16 cases and 16 controls from Cork (<34 weeks of gestation cases plus matched controls) analysed in the untargeted method were included as part of the 162 patients selected for targeted analysis. In addition, 6 of the cases and their 12 matching controls also had plasma collected at time of delivery (TOD). Clinical data and demographics of the studied populations by subset are presented in **Table 5.1**. A flow diagram of the experiment is presented in **Figure 5.1**.

Table 5.1: Demographic and clinical data of the study population.

	Untargeted			Targeted		
	Control (n=39)	**Case (n=39)**	**p-value[b]**	**Control (n=108)**	**Case (n=54)**	**p-value[b]**
Age (years)	30.47 (4.38)	31.28 (5.48)	0.633	31 (19-33)	30 (27-34.25)	0.724
BMI (kg/m^2), grouped according to World Health Organisation						
0 - <18.5, underweight	0	1 (2.6%)		1 (0.9%)	2 (3.7%)	
≥18.5 and <25, normal weight	20 (51.3%)	20 (51.3%)	0.722	65 (60.2%)	33 (61.1%)	0.603
≥25 and <30, overweight	11 (28.2%)	12 (30.8%)		27 (25%)	11 (20.4%)	
≥30, obese	8 (20.5%)	6 (15.4%)		15 (13.9%)	8 (14.8%)	
Ethnicity						
Caucasian	37 (94.9%)	37 (94.9%)		106 (98.1%)	54 (100%)	
Asian	1 (2.6%)	0		0	0	
Pacific Islander	1 (2.6%)	0	0.261	0	0	0.553 [c]
Maori	0	2 (5.1%)		0	0	
African ancestry	0	0		2 (1.9%)	0	
Marital status						
Single	1 (2.6%)	6 (15.4%)		7 (6.5%)	8 (14.8%)	
Married	26 (66.7%)	25 (64.1%)	0.111	74 (68.5%)	32 (59.3%)	0.204
Stable relationship	12 (30.8%)	8 (20.5%)		27 (25%)	14 (25.9%)	
Years of schooling	13 (13-14)	13 (12-13)	0.311	13 (13-14)	14 (13-14)	0.984
Years of schooling, in groups						
<12 years	1 (2.6%)	5 (12.8%)				
12 or 13 years	28 (71.8%)	26 (66.7)	0.227			0.080
>13 years	10 (25.6%)	8 (20.5%)				
Job situation						
Full time work	33 (84.6%)	32 (82.1%)	0.699	88 (81.5%)	43 (79.6%)	0.200

Part time work	3 (7.7%)	2 (5.1%)		12 (11.1%)	3 (5.6%)	
Other	3 (7.7%)	5 (12.9%)		8 (7.4%)	8 (15.0%)	
Socioeconomic index	50 (37-56)	48 (30-56)	0.482	49 (30-60)	44 (27-50)	0.017
Smoking status at 1st visit (15 weeks)						
Never smoked	33 (84.6%)	31 (79.5%)	0.769[c]	80 (74.1%)	39 (72.2%)	0.851[c]
Smoked, past or present	6 (15.4%)	8 (20.5%)		28 (25.9%)	15 (27.8%)	
Alcohol status at 1st visit (15 weeks)						
Did not consume alcohol during pregnancy,	13 (33.3%)	18 (46.1%)	0.355[c]	24 (22.2%)	14 (25.9%)	0.695[c]
Consumed alcohol during pregnancy	26 (66.7%)	21 (53.9%)		84 (77.8%)	40 (74.1%)	
Birthweight (g)	3578.46 (348.29)	1925.0 (1450.0-2100.0)	**<0.001**	3565.0 (3390.0-3762.5)	2515.0 (1987.0-2792.5)	<0.001
Customised birthweight centile [a]	53.79 (26.56)	37.00 (10.20-77.50)	0.086	48.25 (33.25-74.90)	46.45 (15.50-78.13)	0.295
Gestational age at delivery (weeks)	40.15 (1.20)	32.43 (28.86-33.43)	**<0.001**	40.5 (39.86-41.0)	35.15 (33.64-36.43)	<0.001
Other pregnancy complications[d]						
Preeclampsia	NA[e]	1 (2.6%)		NA[e]	2 (3.7%)	
SGA	NA[e]	8 (20.5%)		NA[e]	8 (14.8%)	
LGA	3 (7.7%)	6 (15.4%)	0.247	9 (8.3%)	8 (14.8%)	0.294

Values are shown as mean (SD), median (interquartile range) or n (%).

[a] customised birthweight centile: adjusted for mother's height, weight at 15 weeks of gestation visit, ethnicity, sex and weight of baby and gestation at delivery of baby; [b]Pearson's Chi-squared; [c]corrected by Fisher's exact test; [d]One case had more than one complication; [e]Control group excludes preeclampsia and SGA.

BMI: Body mass index; SGA: small for gestational age; LGA: large for gestational age; NA: not applicable; all mothers were nulliparous.

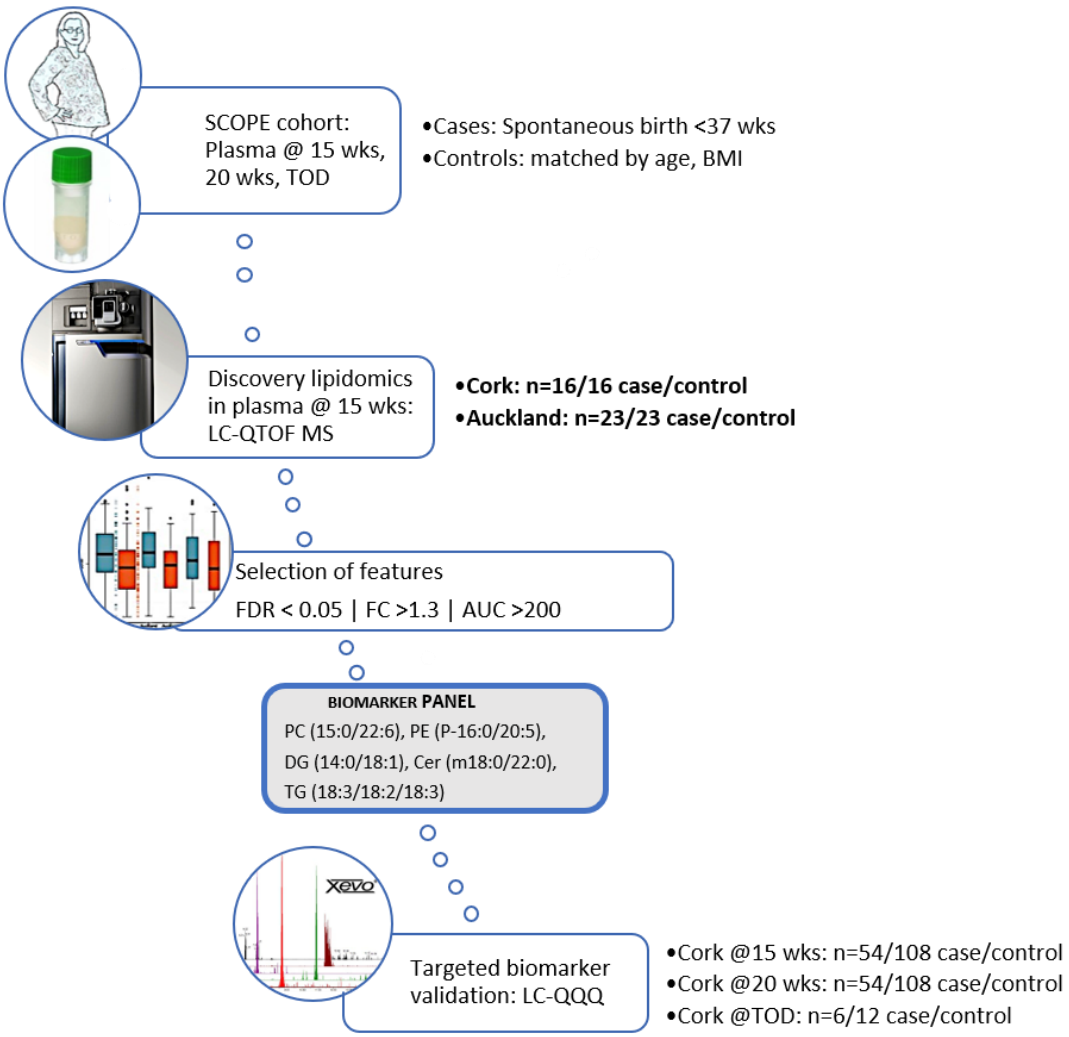


Figure 5.1: Workflow of SCreening fOr Pregnancy Endpoints (SCOPE) participant selection, discovery lipidomics, feature selection and biomarker validation.

5.3.2. Discovery Mass Spectrometry Analysis

Heparinised plasma samples taken at 15 weeks of gestation from Cork and Auckland were prepared and analysed within the same experiment as previously described in detail (Morillon et al. 2020b). In brief, protein precipitation and lipid extraction of 200 µl plasma was performed in a single step by the addition of 600 µl isopropanol (IPA) pre-chilled at -20 °C, based on the methods described by Morillon, *et al.* and Sarafian *et al* (Morillon et al. 2020b; Sarafian et al. 2014). Pooled quality control (QC) samples were created by pooling 30 µl from each sample extract (n=22) vortexed and aliquoted.

For the discovery lipidomic analysis, sample extracts were analysed in a randomised order in triplicate, using ultra performance liquid chromatography quadrupole time-of-flight (UPLC™-QTOF) ACQUITY–SYNAPT G2-S system (Waters Corp., Wilmslow, UK) fitted with an ACQUITY® BEH C18, 1.7 µm, 2.1 x 100 mm analytical column (Waters Corp., Milford, MA, USA). Pooled QC samples were injected to perform column conditioning (n=8 injections) before the start of the analysis and were subsequently injected after every ten samples. An elution gradient of mobiles phases A and B was applied over 23 minutes, at a flow rate of 0.4 ml/min. Mobile phase A was ACN:water (v:v, 60:40) with 10 mM of ammonium formate; mobile phase B was IPA:ACN (v:v, 90:10) with 10 mM of ammonium formate. Mass spectrometry data were collected in continuum format using positive and negative electrospray ionisation (ESI +/-) and MS^E mode (Bateman et al. 2002), a data independent acquisition (DIA) mode. Further details of the chromatography solvent gradient and instrument parameters may be found in the article by Morillon, *et al*. (Morillon et al. 2020b).

5.3.2.1. Data processing and statistical analysis for discovery lipidomic profiling

Progenesis QI version 2.4 (Nonlinear dynamics, Newcastle, UK) was used to process the MS^E data. An appropriate pooled QC was used as the alignment reference to match peaks in each chromatogram according to m/z and retention time. Raw abundance data were normalised to all compounds. Adducts corresponding with

M+H, M+H-H_2O, M+H-2H_2O, M+NH_4, M+Na, M+K, M+2H, M+2Na were considered during peak picking of ESI+ data, and M-H, M-2H, M-H_2O, M+Na-2H, M+Cl, M+K-H were used for ESI-. Compound measurements (raw and normalised compound abundances) were exported from Progenesis QI for downstream statistical analysis.

IBM SPSS Statistics 24 software was used to analyse the demographics and clinical data of the participants, and Student T test, Mann-Whitney U test, Pearson Chi square test, and multiple comparison corrections were applied as appropriate (**Table 5.1**). Block randomisation based on the patient's BMI and the outcome was used for the randomisation of samples before their preparation.

Because of the known differences in analytical sensitivity and technical variability between ESI+ and ESI-, both data sets were analysed independently; statistical analyses were performed using the raw data exports from Progenesis QI with the R statistical software (R Core Team 2013) and the Bioconductor package limma (Ritchie et al. 2015). Median normalisation of the data was performed before applying quality control procedures to check for systematic measurement biases (Broadhurst et al. 2018). Measurement precision was checked for each feature by computing the missing rate and coefficient of variation (CV) over the replicate measurements. Features with a missing rate greater or equal to 20 % and features with a CV greater or equal to 30 % were not considered in subsequent analyses. Features predictive for the outcome were selected using the empirical Bayes (eBayes) method, adjusting the design for replicate measures while identifying differential features (Smyth 2004). This method has been successfully applied to LC-MS data (Kammers et al. 2015; Margolin et al. 2009). This approach was verified by comparing features selected by eBayes to feature selection made using Mann Whitney U test, on the average measurement per patient with multiple testing correction. Agreement between the two methods was demonstrated (Data not shown).

5.3.2.2. Lipid putative annotation

Using the Progenesis QI identification tool, the exact mass of features with a significant adjusted value ($p < 0.03$) were searched against the Human Metabolome

Database version 4.0 (David S. Wishart et al. 2018), LIPID MAPS version of January 2019 (Cotter et al. 2007) and ChemSpider (searching the databases ChEBI version of February 2018, and KEGG version of April 2014; http://www.chemspider.com). Search parameters were set for an exact mass tolerance of 5 ppm (precursor ions) and 10 ppm (fragment ions). Only putatively annotated compounds with a significant difference between cases and controls (eBayes corrected p-value <0.05) were included in the initial short list of potential features. Five features from this short list were selected for the targeted analysis.

5.3.2.3. Selection of biomarker candidates for targeted validation

To select putative biomarkers for validation by multiple reaction monitoring (MRM) with the XEVO™ TQ MS, stringent filtering criteria were applied to the discovery data including a fold change (FC) $\geq$1.3, Area Under the Curve (AUC) of >200 counts and eBayes corrected $p < 0.05$. Verification was performed by matching retention times of the selected features from the SYNAPT with those of the XEVO™ TQ MS using the same column and chromatography conditions in positive or negative mode as appropriate. As authentic standards were not available, the instrument was tuned using the SPLASH® Lipidomix®, which contains deuterium labelled lipids representative of the major lipid classes. The SPLASH® Lipidomix® does not contain a ceramide, instead a purchased ceramide Cer (m18:0/24:1) standard was used to tune the cone voltage and collision energy for this class. Three more candidates were eliminated as targets due to poor chromatographic peak shape, low AUC on the XEVO™ TQ MS (<200), or mismatch between retention times and fragmentation patterns of the XEVO and SYNAPT. MRM methods were created to determine the specific fatty acyl groups associated with each lipid.

5.3.3. Targeted mass spectrometry analysis

5.3.3.1. Sample preparation

All sample aliquots (250 µL) were thawed and 10 µL from each 15 weeks of gestation sample and 20 µL from each 20 weeks of gestation sample were pooled together to create quality control (QC) samples at 15 and 20 weeks of gestation (QC15 and QC20, respectively). Sample IDs were randomly blocked into 7 batches, with each participant's time point samples run consecutively, and each batch was thawed and extracted in a single day. The standard curve was prepared by pipetting 50 µL each of 7 standards in methanol prepared from dilutions of SPLASH® Lipidomix® (**Appendix D**). The methanol was evaporated in a vacuum centrifuge, and 50 µL of plasma matrix (pooled plasma from healthy, non-pregnant females) was added into each tube and vortexed for 30 seconds.

50 µL plasma of each sample was pipetted into labelled Eppendorf tubes, and QC15 and QC20 were prepared in triplicate for every batch. Chilled IPA was added to each tube (standards and samples) and vortexed 20 seconds. Samples were kept at -20 °C for 10 minutes, then at 4 °C for 2 hours. The samples were centrifuged for 10 minutes at 14,000 g, and 200 µL supernatant transferred to HPLC vials for analysis. The standard curve was run at the beginning and the end of each batch with QC samples acquired every 12th injection. The injection volume was 10 µL using full loop mode, and autosampler was set to 4 °C.

5.3.3.2. LC-MS/MS: targeted method

For the targeted LC-MS/MS analysis we employed a Waters ACQUITY-XEVO™ TQ MS triple quadrupole (TQ) instrument (Waters Corp., Wilmslow, UK). The mobile phases A and B and analytical column were identical to that used in the untargeted analysis. The weak needle wash was 50:50 ACN:water, while the strong needle wash was 90:10 IPA:ACN. The gradient was modified to shorten the run time from 23 minutes to 14 minutes as follows: 30 % B initial, gradient to 70 % B for 10 minute, gradient to 91.3 % B for 1.5 minutes, hold at 99 % B for 1 minute, return to initial for 1.5 minutes, end time 14 minutes. The XEVO™ TQ MS was operated in electrospray positive mode

with a capillary voltage set to 2.7 kV, extractor to 3.0 V, desolvation temperature at 500 °C, cone gas flow at 50 L/h, desolvation gas flow at 1000 L/h, collision gas flow at 0.15 mL/h. Argon was the collision gas, while the desolvation gas was nitrogen. The cone voltages and collision energies of each target and labelled analytes from the SPLASH® Lipidomix® are shown in **Table 5.2**, and their corresponding chromatograms are shown in Fig 2. Results were processed using TargetLynx™ software Version 4.1 (Waters Corp., Wimslow, UK).

Semi-quantification of the samples was achieved using Average Response Factor in TargetLynx™ of the nominally assigned concentration of the target lipids naturally present in the calibrator matrix (**Appendix D** Method for further details). Analytes were quantified using peak area, with the exception of TG (18:3/18:2/18:3) and TG (15:0/18:1 (d7)/15:0) which were quantified using peak height because of incomplete peak resolution (see inset, **Figure 5.2**). The linear calibration curves were used for quality control purposes, with batch acceptance criteria of $\pm$30 % internal standard recovery of 4 out of 6 QC samples within each batch (**Appendix D**).

Table 5.2: Retention times and Multiple Reaction Monitoring (MRM) parameters for five selected lipid targets and four deuterium labelled lipid standards.

Lipid	RT (min)	Neutral Mass (Da)	Transition	Adduct	Cone (V)	Coll (eV)
PC (15:0/22:6)	6.32	791.55	792.6 > 184.1	$[M+H]^+$	40	25
			792.6 > 482.3	$[M+H]^+$	40	25
PC (15:0-18:1-d_7)	7.52	753.11	753.6 > 184.1	$[M+H]^+$	42	10
PE (P-16:0/20:5)	7.08	721.50	722.5 > 581.3	$[M+H]^+$	48	34
PE (15:0/18:1-d_7)	7.74	711.03	711.7 > 288.4	$[M+H]^+$	48	34
DG (14:0/18:1)	8.68	566.49	584.5 > 285.2	$[M+NH_4]^+$	18	20
			584.5 > 339.3	$[M+NH_4]^+$	18	20
DG (15:0/18:1-d_7)	9.09	587.98	605.6 > 346.4	$[M+NH_4]^+$	18	20
Cer (m18:0/22:0)	10.9	607.63	608.6 > 268.3	$[M+H]^+$	40	26
			608.6 > 590.6	$[M+H]^+$	40	20
TG (18:3/18:2/18:3)	11.46	874.71	892.8 > 595.2	$[M+NH_4]^+$	42	24
			892.8 > 597.5	$[M+NH_4]^+$	42	24
TG (15:0/18:1-d_7/15:0)	12.92	812.37	829.9 > 523.5	$[M+NH_4]^+$	38	24

PC = phosphatidylcholine; PE = phosphatidylethanolamine; DG = diacylglycerol; Cer = ceramide; TG = triacylglycerol.

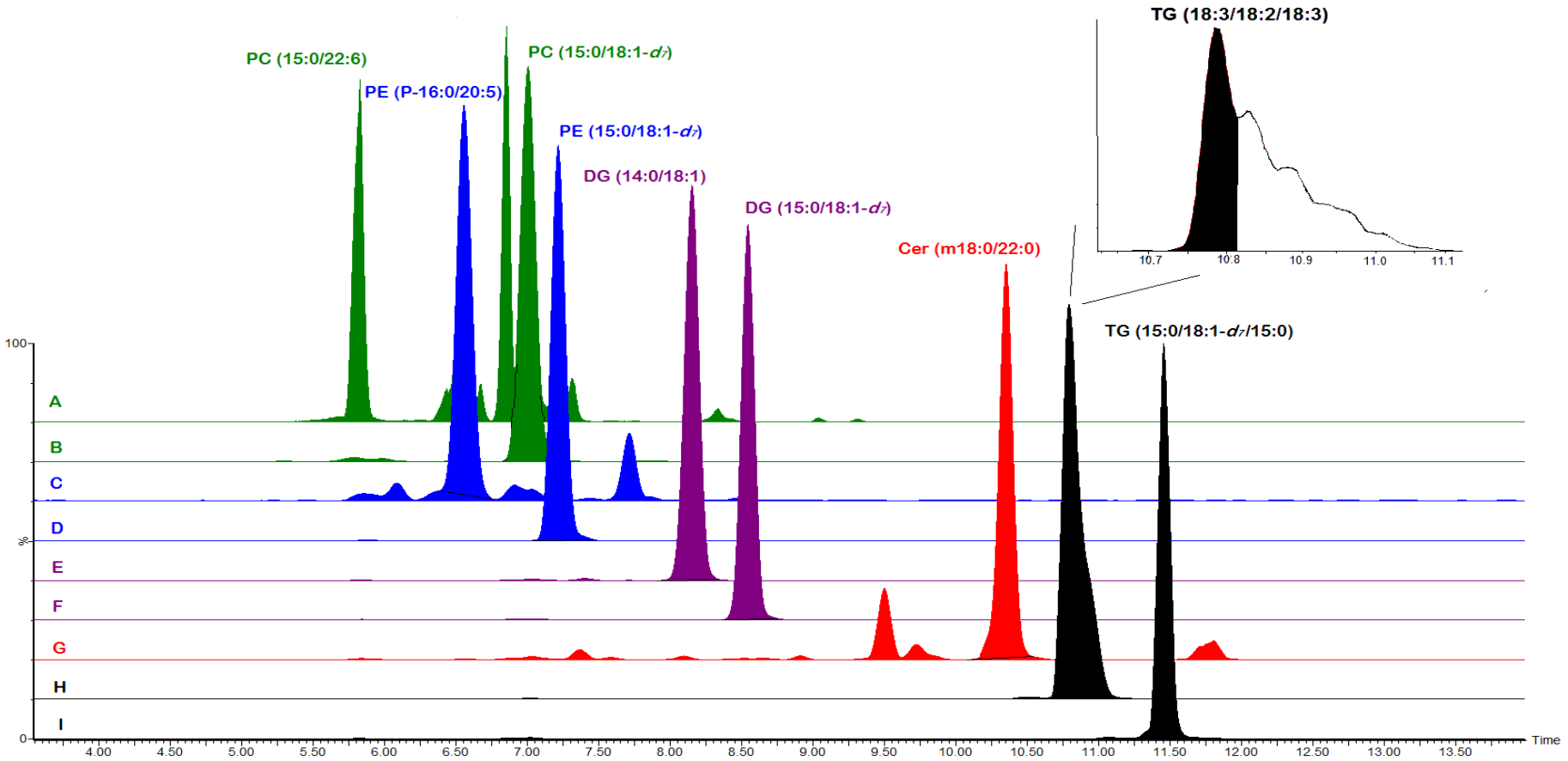


Figure 5.2: Chromatograms of a representative plasma sample. A: Phosphatidylcholine (PC) (15:0/22:6), B: PC (15:0/18:1-d7), C: Phosphatidylethanolamine (PE) (P-16:0/20:5), D: PE (15:0/18:1-d7), E: Diacylglycerol (DG) (14:0/18:1), F: DG (15:0/18:1-d7), G: Ceramide (Cer) (m18:0/22:0) H: Triacylglycerol TG (18/18:2/18:3), I: TG (15:0/18:1-d7/15:0)

The relative concentrations of the targeted lipids in pooled QC samples at 15 and 20 weeks of gestation and their intra- and inter-day CVs are shown in **Appendix D**. The intra- and inter-day CV was less than 19 % and 10 %, respectively, for all analytes except the Ceramide which was >39 % at both time points. The relative concentrations of DG (14:0/18:1) and TG (18:3/18:2/18:3) were higher in the 20 weeks of gestation QC compared to 15 weeks of gestation QC, whereas there was no difference in the concentration of PC (15:0/22:6) and PE (P-16:0/20:5) between the two time points.

5.3.3.3. Statistical Analysis for targeted validation of lipid markers

Statistical comparisons of biomarker concentrations in sPTB cases vs. age and BMI matched controls were produced using SPSS v.25 (IBM, Inc.). Microsoft Excel 2016 was used to calculate method performance parameters of the analytes and produce graphical output of statistical comparisons. Variables were checked for normal distribution using the Kolmogorov-Smirnoff test; non-normally distributed variables were natural log transformed prior to subsequent statistical tests. All variables had equally distributed variances according to Levene's test (>0.05). Independent t-tests compared biomarker concentrations between case vs. control groups at 15 and 20 weeks of gestation. Random forest model using the targeted lipids on their own, or in combination with clinical data were built at 15 and 20 weeks of gestation time points. The clinical data used to build this model included age, gravidity, BMI, marital status, socioeconomic index, education level, maternal birthweight, being a current smoker, and being a current alcohol consumer. Significance was defined as $p < 0.05$.

5.4. Results

We undertook a discovery lipidomics study using plasma samples collected at 15 weeks of gestation from the SCOPE cohort in Cork and Auckland in order to identify lipids that are differentially expressed in women who had sPTB compared to women with uncomplicated, full-term delivery. Of the 4057 compounds detected in negative

mode and 6438 compounds detected in positive mode, empirical Bayes statistical methods combined with chemical database searching identified 122 differentially expressed lipids ($p < 0.05$) with a putatively assigned chemical name and lipid class (**Appendix D**), representing five lipid classes (**Table 5.3**). While the identities of the majority of these lipids were not positively confirmed, a general trend of decreased glycerophospholipids (80 % of glycerophospholipids) in sPTB cases was observed.

Table 5.3: Summary of putatively annotated differential features in spontaneous preterm birth cases relative to full term birth controls. eBayes adjusted for FDR at 5 % p-value <0.05.

Chemical class	Number of lipids identified*	Lipids selected for targeted analysis	Change relative to control	Normalised abundance Fold Change**	eBayes p value**
Glycerolipid	30		Up (n=14); Down (n=16)		
		DG(32:1)	Up	1.33	5.8E-03
		TG(54:8)	Up	1.48	5.3E-03
Glycerophospholipid	77		Up (n=15); Down (n=62)		
		PC(37:6)	Down	1.63	1.2E-02
		PE(P-16:0/20:5)	Down	1.33	7.4E-04
Sphingolipid	12		Up (n=7); Down (n=5)		
		Cer(m18:0/22:0)	Up	1.31	8.5E-03
Sterol lipid	3		Up (n=2); Down (n=1)		

*Lipids putatively annotated using Progenesis software resulting from Human Metabolome Database and LIPIDMAPS searches, **Statistic provided by SQU4RE. Bold: Features selected for targeted analysis. DG = Diacylglycerol; TG = Triacylglycerol; PC = phosphatidylcholine; PE = phosphatidylethanolamine; Cer = Ceramide.

5.4.1. Selection of biomarker candidates for targeted validation

Using the criteria outlined in methods, we identified five putative biomarkers to take forward for targeted validation by MRM including PC (15:0/22:6), PE (P-16:0/20:5), DG (14:0/18:1), Cer (m18:0/22:0) and TG (18:3/18:2/18:3) (**Table 5.3**). Comparisons of sPTB cases (n=39) to controls (n=39) at 15 weeks of gestation of the five compounds selected from the discovery lipidomic phase of the study are shown in **Figure 5.3A-E**. PC and PE were significantly downregulated relative to control (FC = 1.63 and 1.33, respectively) whilst DG, Cer and TG were significantly upregulated (FC = 1.33, 1.48 and 1.31, respectively).

5.4.2. Targeted analysis and validation of biomarker panel

Targeted biomarker validation was performed on n=54 cases and n=108 controls from the SCOPE cohort in Cork at 15 and 20 weeks of gestation. Results for targeted MRM analysis of PC (15:0/22:6), PE (P-16:0/20:5), DG (14:0/18:1), Cer (m18:0/22:0) and TG (18:3/18:2/18:3) are shown in **Figure 5.3F-I**. Statistical comparison of each individual compound in cases vs. controls at each time point (independent t-test) showed that PC (15:0/22:6) was detected in significantly lower levels in cases compared to controls at 15 weeks of gestation (p=0.015, FC=1.12; **Figure 5.3F**), confirming the trend observed in the discovery analysis (p= 0.012 , FC = 1.63; **Figure 5.3A**).

TG (18:3/18:2/18:3) was significantly increased in sPTB cases compared to controls at 20 weeks of gestation (p=0.002, FC=1.63, **Figure 5.3I**), and although a similar increased trend was observed in the same women at 15 weeks of gestation, the difference was not significant (p-value, fold change), as predicted in the discovery analysis in Cork and Auckland at 15 weeks of gestation (p=0.005, FC=1.48; **Figure 5.3D**).

PE (P-16:0/20:5) and DG (14:0/18:1) showed no significant differences between cases and controls at 15 or 20 weeks of gestation (**Figure 5.3G,H**), suggesting the findings observed in the discovery analysis were false positives or perhaps attributed to lipid

alterations in the Auckland cohort of samples. The ceramide was eliminated as a target because the MS signal decreased progressively from Day 1 to Day 8 resulting in an inter-day CV >39 %, indicating that this analyte was not stable (**Appendix D**), and the results presented in **Figure 5.3D** are a likely false positive.

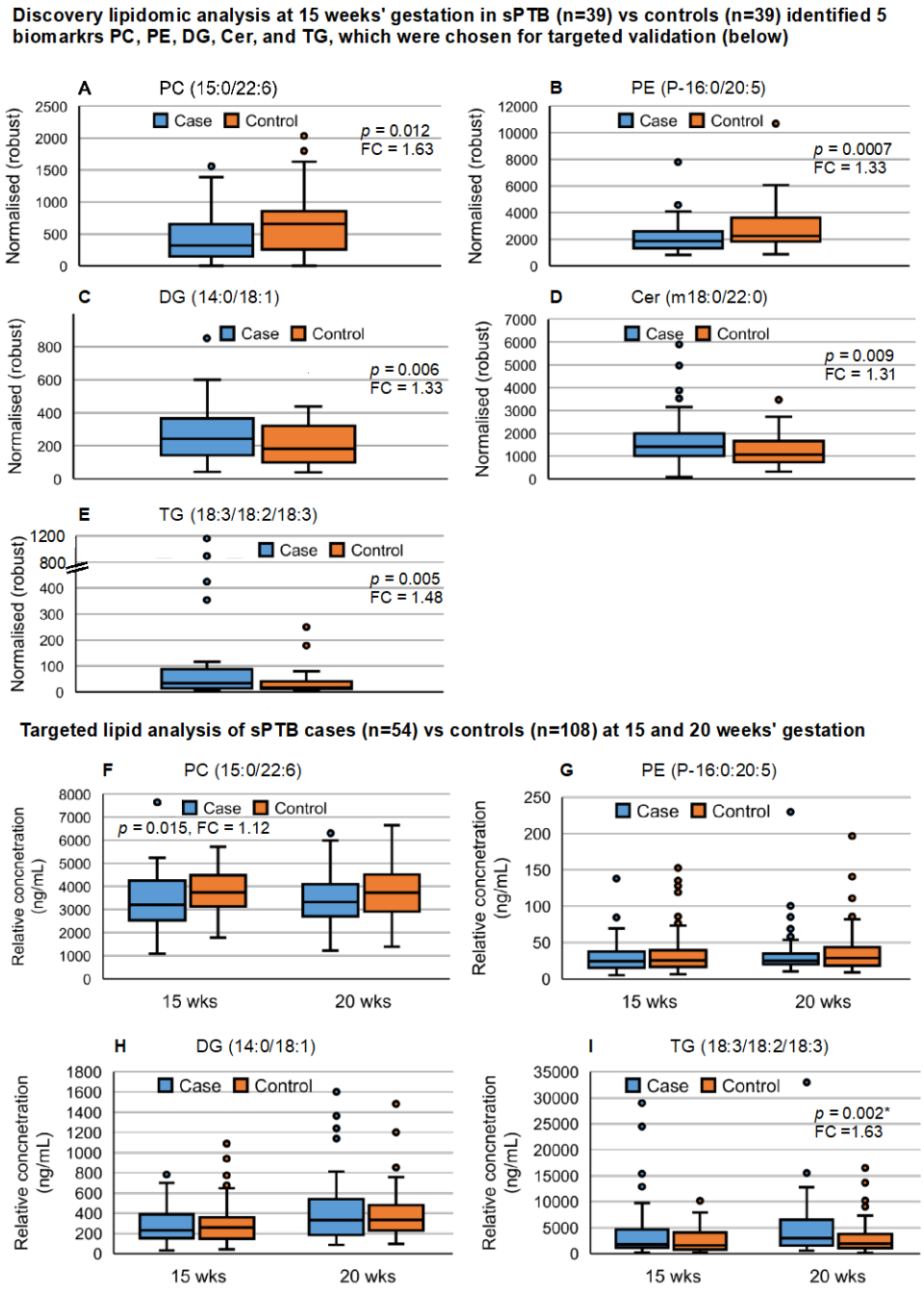


Figure 5.3: Box plots showing the normalised intensity of the selected lipids significantly altered (corrected p <0.05) between cases and controls at 15 weeks,

from untargeted UPLC-Q-TOF-MS screen of SCOPE Cork and Auckland plasma samples (case n= 39, control n=39). A: Phosphatidylcholine (PC) (15:0/22:6), B: Phosphatidylethanolamine (PE) (P-16:0/20:5), C: Diacylglycerol (DG) (14:0/18:1), D: Ceramide (Cer) (m18:0/22:0), E: Triacylglycerol (TG) (18:3/18:2/18:3); and box plots showing the semi-quantified targeted analysis of the lipids in 54 cases and 108 controls at 15 weeks and 20 weeks of gestation F: PC (15:0/22:6), G. PE (P-16:0/20:5) H. DG (14:0/18:1) I. TG (18:3/18:2/18:3). Cer (m18:0/22:0) failed repeatability criteria (inter-day CV >30 %); targeted results not shown. Outliers (circles) are greater than 1.5 times upper quartile. FC: Fold change.

In addition, 6 cases and 12 matched controls were collected at the time of delivery (TOD) and analysed. Statistical analysis comparing cases to controls at 15 and 20 weeks of gestation and time of delivery shows a trend, where TG (18:3/18:2/18:3), PC, and PE levels seem to peak between 15 and 20 weeks of gestation and decrease at time of delivery. A parallel coordinate plot of all four targeted lipids illustrates their levels between sPTB cases and controls over time (**Figure 5.4**).

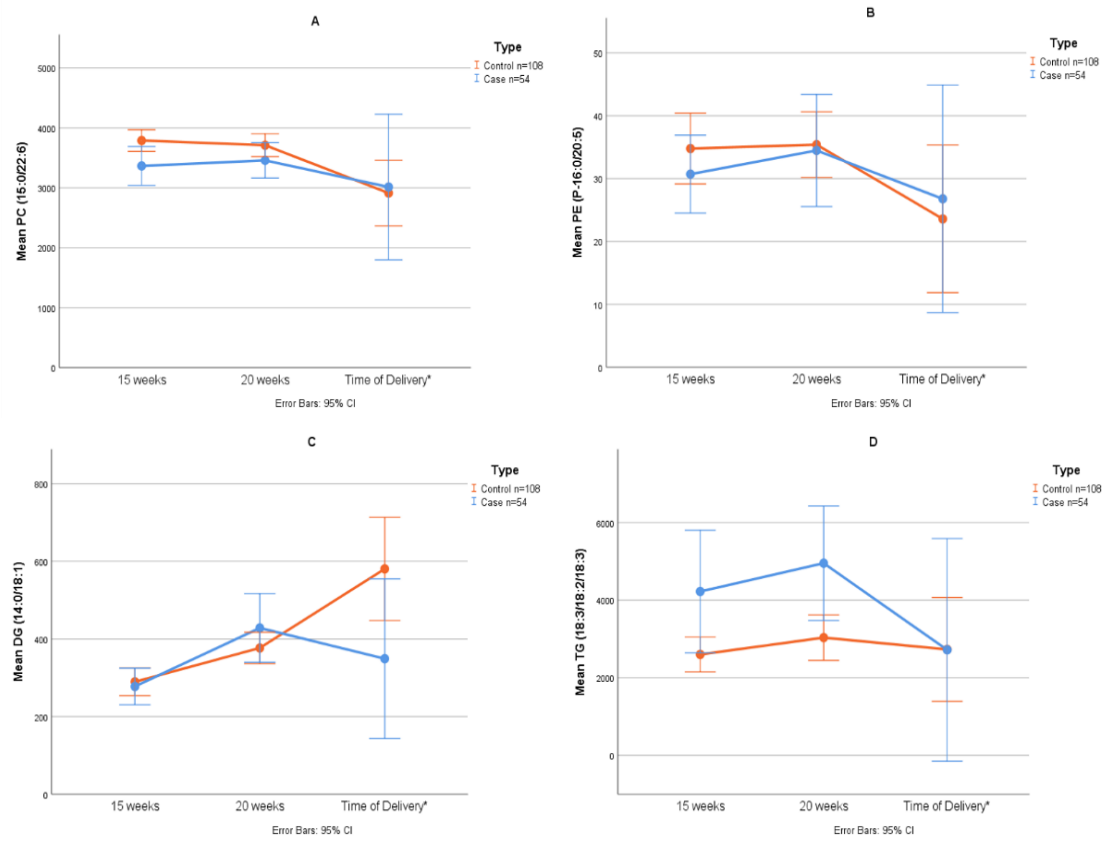


Figure 5.4: Targeted MRM analysis of PC, PE, DG and TG lipids in maternal plasma samples collected at 15 weeks and 20 weeks of gestation, and at time of delivery (TOD). Parallel coordinate plots are used to illustrate the mean intensity (95 % confidence interval [CI]) of the four targets. Blue lines represent the sPTB cases (n=54), and red lines represent their matched controls (n=108). *For the TOD time point, just 6 cases and 12 control samples were available for analysis. A: Phosphatidylcholine (PC) (15:0/22:6), B: Phosphatidylethanolamine (PE) (P-16:0/20:5), C: Diacylglycerol (DG) (14:0/18:1), D: Triacylglycerol (TG) (18:3/18:2/18:3).

Finally, receiver operating characteristic (ROC) curves were used to test the predictive value of the four targeted lipids ((PC (15:0/22:6), PE (P-16:0/20:5), DG (14:0/18:1), Cer (m18:0/22:0), TG (18:3/18:2/18:3)) alone or in combination with the clinical data model at 15 or 20 weeks of gestation are presented in **Figure 5.5**. The AUC of the four lipids targeted at 15 weeks of gestation is 0.57 (**Figure 5.5A**), while it is 0.56 at 20 weeks of gestation (**Figure 5.5C**). Combining these markers with the clinical data did not improve their prediction value, and the mean AUC at 15 weeks of gestation is 0.50 (**Figure 5.5B**), and 0.49 at 20 weeks of gestation (**Figure 5.5D**).

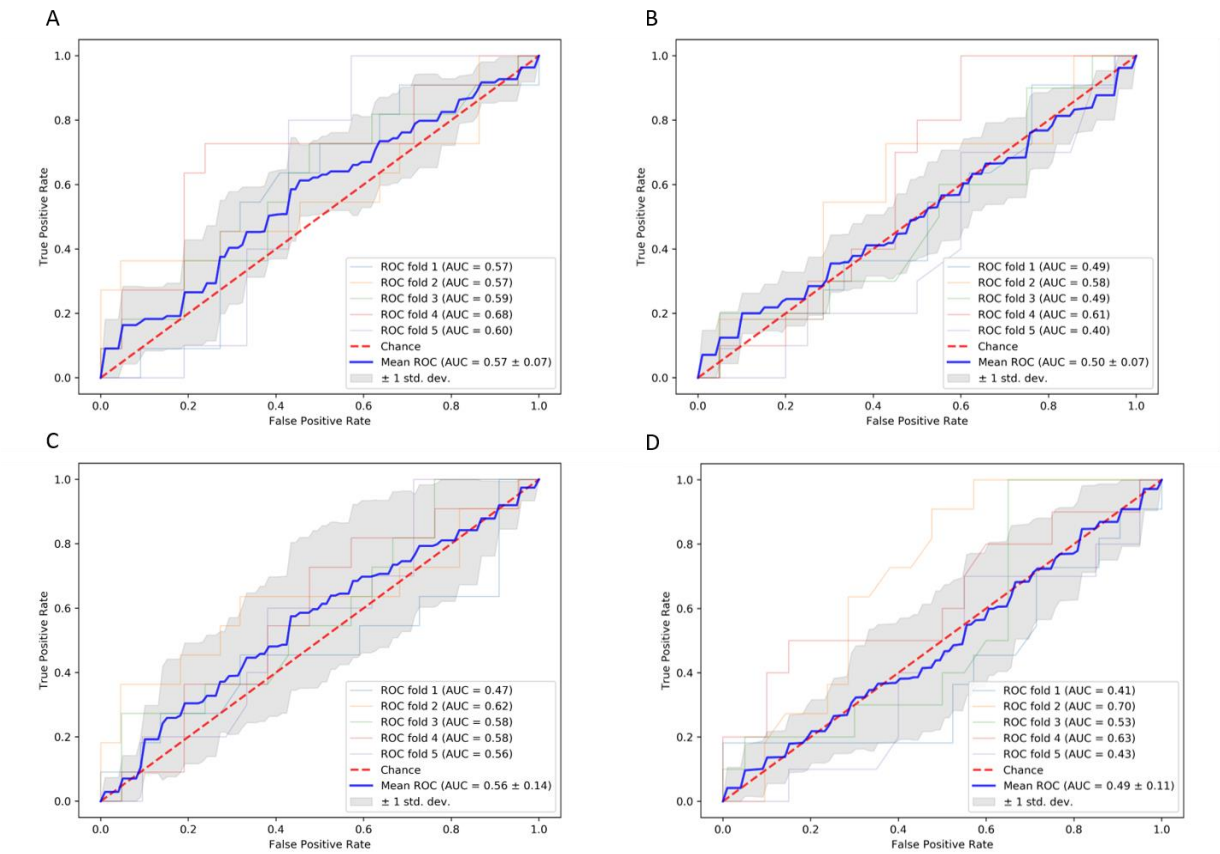


Figure 5.5: Prediction model for sPTB <37 weeks of gestation, using the four targeted lipids PC (15:0/22:6), PE (P-16:0/20:5), DG (14:0/18:1), TG (18:3/18:2/18:3) alone or combined with clinical data. Folds represent sequential iterations of the test 5-fold cross validation applied for the generation of the training and test datasets, on which the models were constructed and assessed to reduce the risk of overfitting. The mean value for AUC was calculated from the overall proportions of correct predictions from all folds. A: Four targeted lipids at 15 weeks of gestation; B: Four targeted lipids and clinical data at 15 weeks of gestation; C: Four targeted lipids at 20 weeks of gestation; D: Four targeted lipids and clinical data at 20 weeks of gestation.

5.5. Discussion

Our findings indicate that the lipidome is altered in plasma between 15 and 20 weeks of gestation in women who subsequently had sPTB. These results give promise to the potential of an early biomarker signature for sPTB. At 15 weeks of gestation, our discovery analysis in Cork and Auckland SCOPE Cohorts found 122 lipids to be significantly altered, with 77 phospholipids found to be significantly reduced in the sPTB group in comparison to controls (39 cases vs 39 controls). We previously reported reduced phospholipid levels at 20 weeks of gestation in the Cork cohort samples (Morillon et al. 2020b), and together these findings are indicative of altered phospholipid metabolism in the pathophysiology of sPTB. We further extended our findings in a targeted validation of 5 putative lipid biomarkers at 15 and 20 weeks of gestation in an extended sample (n=54 cases; 108 controls) from Cork SCOPE (www.scopestudy.net). We confirmed a reduction in PC (15:0/22:6) at 15 weeks of gestation in sPTB, and an increase in TG (8:3/18:2/18:3) at 20 weeks of gestation in sPTB.

This is the first lipidomic study reporting both discovery and targeted validation of potential lipid biomarkers for prediction of sPTB. We previously showed the potential of lipidomics to characterise the plasma profile at 20 weeks of gestation in sPTB in the Cork SCOPE cohort (Morillon et al. 2020b), whereby we reported an overall reduction in 26 phospholipids, including glycerophospholipids and sphingolipids in sPTB in comparison to controls. In this study, we expand on these findings to implicate a similar trend in altered phospholipids at 15 weeks of gestation in both Cork and Auckland populations of the SCOPE cohort, suggesting phospholipid metabolism maybe broadly implicated pathophysiology of sPTB.

Glycerophospholipids (a phospholipid subgroup) are involved with various cellular processes, including storage and breakdown of energy, inflammation, apoptosis, and cell-membrane stabilisation (Baig et al. 2013). Glycerophospholipids, such as sphingolipids and phosphatidylcholines, have been associated with adverse pregnancy outcomes in animal model of pregnancy loss (Mizugishi et al. 2007), and in human studies of small for gestational age (Horgan et al. 2011), preeclampsia and recurrent pregnancy loss (Baig et al. 2013). The most abundant phospholipid in

human blood samples are Phosphatidylcholines (PC) (Uran et al. 2001; Schwarz et al. 1977), which are crucial for cellular membrane integrity and function (Castro-Gómez et al. 2015). According to Schwarz and colleagues, the human blood plasma concentrations of phospholipids is 1450 µM of PC, 473 µM of SM, 175 µM of PE, 114 µM of PI, and 63 µM of PG (Schwarz et al. 1977). Taken together, phospholipids in blood are estimated to have a concentration of 2900 µM (Schwarz et al. 1977). Other species of metabolites present in human blood samples include amino acids, detected in concentrations ranging from 1 µM to approximately 1200 µM (Gray et al. 2017), which are present in similar dynamic range as phospholipids. In our discovery analysis we found reduced PC (15:0/22:6), FC=-1.63 at 15 weeks of gestation in Cork and Auckland participants (39 cases vs 39 controls) from SCOPE. We subsequently targeted and confirmed this finding in an extended sample from Cork participants (54 cases vs 108 controls) at 15 weeks of gestation. In addition, we profiled the expression of this marker in the same women at 20 weeks of gestation (**Figure 5.4A**) and time of delivery (TOD), and found the differential expression of PC is neutralised at 20 weeks of gestation, with a downward trend in expression in both sPTB cases and controls at TOD. It is interesting to note that studies have shown that in uncomplicated pregnancies, the maternal plasma levels of PC increased between the second (16 and 22 weeks) and third (32 and 40 weeks) trimesters (Postle et al. 1995). (West et al. 2013; Yan et al. 2013). During the third trimester, PC enriched in species containing arachidonic acid (20:4) and docosahexanoic acid (22:6), which are synthesised through the hepatic maternal phosphatidylethanolamine-*N*-methyltransferase (PEMT) pathway, are selectively transported to the fetus (Yan et al. 2013). The alternative pathway of PC synthesis is through choline incorporation into PC via the cytidine diphosphocholine (CPD-choline) pathway, which leads to mostly monounsaturated and di-unsaturated molecular species (Weiss et al. 1958). A recent study using stable isotope labelling of methyl-D_9-choline combined with ESI-MS/MS monitored and quantified the appearance of label PC in plasma. This enabled them to characterise the metabolism and turnover of PC synthesis metabolism in the immature extremely preterm infants, born at less than 28 weeks of gestation and by extension, the fetal liver (Goss et al. 2020). Goss and colleagues demonstrated that the fetal liver showed low PEMT activity and high activity of the hepatic CDP-choline

pathway (Goss et al. 2020). Taken together with the results presented in this chapter, this shows that the significantly lower levels of PC (15:0/22:6) at 15 weeks, but normalised levels at 20 weeks in sPTB group compared to control group, could reflect alterations of the transport of PC(15:0/22:6) to the fetus. However, more experiments are needed to further understand these alterations.

In addition, a deficiency in enzymes responsible for PC biosynthesis in KO mouse models has been associated with several pathological phenotypes (van der Veen et al. 2017), including embryonic lethality (Gengshu Wu et al. 2008; L. Wang et al. 2005), neonatal respiratory failure (Tian et al. 2007), impaired neuronal branching (Carter et al. 2008), bone deformity (Z. Li et al. 2014), and muscular dystrophy (Sher et al. 2006). The same deficiency in human has been associated with muscular dystrophy (Mitsuhashi et al. 2011), lipodystrophy and non-alcoholic fatty liver disease (NAFLD) (Payne et al. 2014). Moreover, a recent study showed that the risk of contracting Alzheimer's disease was reduced when taking supplementary dietary PC-DHA (Patrick 2019). And PC levels in human plasma samples have also been linked with risk of cardio vascular disease, via the trimethylamine-N-oxide biosynthesis in the intestinal microbiota (W. H. W. Tang et al. 2013).

Although only 6 cases and 12 control samples were analysed at TOD, our findings highlight the benefit of longitudinal targeted profiling of putative markers during pregnancy, to capture the plasma molecular signatures at play.

Triacylglycerols (TG) are synthesised by intestinal and liver cells and enclosed in lipoproteins for circulation in the bloodstream, transporting TG to peripheral tissues where they may be hydrolysed to release energy or stored in adipose tissue (Yen et al. 2008). Elevated triacylglycerols in the circulation increase the risk of heart disease and stroke, and have previously been associated with preterm delivery (Catov et al. 2007; Zheng et al. 2018) and adverse pregnancy outcomes (Nasioudis et al. 2019). Total plasma TG have been shown to increase throughout gestation in healthy pregnancy, with TG peaking right before delivery (dos Santos-Weiss et al. 2013; Hubel et al. 1998; Sattar et al. 1997). Here we found TG (18:3/18:2/18:3) markedly increased in sPTB in comparison to controls (39 cases vs 39 controls) at 15 weeks of gestation in the Cork and Auckland cohorts ($p=0.002$, FC=+1.63). Our targeted

analysis in an extended sample of from Cork participants (54 cases vs 108 controls) did not confirm significance at 15 weeks of gestation (p=0.11), but an increased trend was evident (FC=+1.56), and a significant increase was observed at 20 weeks of gestation (p=0.002, FC=+1.63) in the sPTB group. Unlike total TG which peaks before delivery, TG (18:3/18:2/18:3) decreased at time of delivery in both cases and controls (**Figure 5.4**) and additional samples taken at time of delivery would be needed to confirm this downward trend.

It is noteworthy that the longitudinal profile of PC, PE and TG markers follow a similar expression profile at 15, 20 weeks of gestation and TOD (**Figure 5.4**) in cases and controls during pregnancy and may represent overall physiological plasma profile changes during pregnancy. While the difference between DG (14:0/18:1) in cases and controls was not statistically significant (p= 0.8), a trend towards increased DG at TOD in controls but not cases was observed. Diacylglycerols increase at parturition as the oxytocin/oxytocin receptor pathway is activated (McCormick 1985). Further investigation of this pathway at TOD may show alterations in preterm initiation of labour.

The acyl chains of TG (18:3/18:2/18:3) and PC (15:0/22:6) are comprised of Polyunsaturated Fatty Acids (PUFA) which may be either from the omega-3 or omega-6 fatty acid series, both of which are obtained through dietary intake. While our findings did not specifically implicate circulating PUFAs, cause and effect of altered PC (15:0/22:6) and TG (18:3/18:2/18:3) containing PUFAs warrant investigation as potential therapies to delay preterm birth. Omega-3 FAs (e.g. alpha-Linolenic acid - ALA, eicosapentaenoic acid EPA, and docosahexaenoic acid- DHA) have been shown to regulate plasma lipid levels (Rambjør et al. 1996; Mori et al. 2000), and were altered in placental tissue of infants delivered preterm compared to full term births (Dhobale et al. 2011). In their rat model of pregnancy, Chalil *et al.* (Chalil et al. 2018) showed that DHA levels in plasma PCs at day 20 of pregnancy (12.67±1.13 mg/100 ml) were significantly increased compared to 15 days of pregnancy (6.89±1.45 mg/100 ml), and baseline (prior to pregnancy, 5.69±0.86 mg/100 ml). In addition, their study showed that FA composition of plasma TG containing ALA was significantly increased at 15 days of pregnancy (0.92±0.29

mg/100 ml) compared to baseline (0.89±0.26 mg/100 ml), and significantly higher than at 20 days of pregnancy (0.57±0.19 mg/100 ml). Importantly, a recent meta-analysis of 70 randomised control trials assessing the effect of omega-3 FA supplementation on preterm birth showed that omega-3 FA supplementation (specifically, DHA and EPA) reduced the risk of early preterm birth (<34 weeks of gestation) by 58 % (relative risk: 0.42; 95 % confidence interval 0.27-0.66; p=0.0002), and the risk of all preterm delivery was reduced by 17 % (relative risk: 0.83; 95 % confidence interval 0.70-0.98; p=0.03) (Kar et al. 2016). Together, these findings emphasize the potential of omega-3 supplementation as an effective and simple strategy to reduce preterm birth at a population level.

Our study has its limitations. The four lipids selected for validation in an extended sample size from the Cork SCOPE cohort (54 cases vs 108 controls) did not show adequate prediction as potential biomarkers of sPTB at 15 or 20 weeks of gestation (AUC 0.57 and 0.56, respectively; **Figure 5.5**), in spite of the observed significant alterations at the varying timepoints during pregnancy. Secondly, we selected lipids for validation based on p-value, fold change, and MS signal intensity but alternative methods to select predictive biomarkers using machine learning and artificial intelligence are increasingly being used (Aung et al. 2019; Bahado-Singh et al. 2019) and are likely to result in more efficient selection of predictive biomarkers. A third limitation of our study was the lack of validation in the Auckland cohort, which may explain the loss of significance of PE and DG markers in our targeted validation and warrants further investigation. Another limitation of the study was the number of samples analysed. Metabolomics expression studies in which the outcome is expected to be a multi-factorial test (multivariate model) are difficult to power, as it is difficult to predict beforehand the combinatorial discriminatory ability of the independently measured factors. However, for the multivariate approach to be successful it is expected that most, if not all, of the influential factors in the optimal model to be significant, with a critical p-value of 0.05, when subjected to univariate testing (e.g. Student's T-test). Assuming a Type I error rate, α, of 5% is sufficient and that will be at least a 50% change in mean (with a similar within sample standard deviation), then in order to test hypotheses to a suitable level of power (95% power,

β = 0.05) the samples size should approximate to 20 cases, calculated using MEDCALC® (www.medcalc.com). We estimate that: given the univariate power calculations (based on similar clinical studies (Menon et al. 2014; Auray-Blais et al. 2011); the strict collection/processing protocols; the close matching of cases and controls; and the relatively homogeneous target population, then for a metabolomics study the number of samples analysed (39 cases vs 39 controls for the discovery analysis; two time points for 54 cases vs 108 controls for targeted analysis) is sufficient for the aims of this study. Finally, despite rigorous selection of the participants used in this study, it is possible that the sPTB group presents several subgroups of sPTB. Indeed, the main selection criteria was that the women selected for the sPTB group had spontaneous premature labour before 37 weeks of gestation. However, the gestational age at birth ranged from 28.86 to 33.43 weeks for the discovery analysis, and from 33.64 to 36.43 weeks in the targeted analysis (**Table 5.1**). This means that the number of weeks of gestations left between the sample collections (15 or 20 weeks) and premature labour was not the same for all participants and could introduce heterogeneity in an already heterogeneous pregnancy complication.

Nonetheless, our discovery lipidomic analysis in geographically distinct pregnancy cohorts, and longitudinal validation of our four putative markers at 15, 20 weeks of gestation and TOD yielded novel insights into the physiological changes that occur over the course of pregnancy. Future biomarker investigations of "At Risk" pregnancy states would benefit in longitudinal molecular profiling during pregnancy, instead of focusing on a single time point. Such a strategy would facilitate the discovery of dysregulated molecular pathways and biological processes that serve as robust predictors of 'At Risk' pregnancy states, thereby facilitating patient stratification and a better understanding of physiological changes that over the course pregnancy.

In summary, there is firm evidence that plasma physiological changes and altered phospholipid metabolism occur at 15 and 20 weeks of gestation (Morillon et al. 2020b) prior to the onset of sPTB, and our results provide valuable insights into the pathophysiology of sPTB. This study highlights the importance of furthering our understanding of the normal physiological changes that occur in pregnancy to help

differentiate from the molecular changes that are abnormal. Future validation work, employing machine learning to select more robust predictors and incorporate clinical risk factors (e.g. fetal, maternal, uterine disorders (Purisch et al. 2008), short cervical length (<25 mm) (Committee on Practice Bulletins 2012), maternal infections (Haahr et al. 2016; Cunnington et al. 2013)) will help with the selection of putative lipids to take forward for biomarker validation, and translate the development of biomarker panels for the early identification and clinical intervention in 'At Risk' pregnancy states.

Chapter 6

Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia.

Morillon, A.-C., et al. (2020). "Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia." PLoS One 15(3): e0230977. (Morillon et al. 2020a)

6.1. Abstract

Introduction

Pre-eclampsia is a major cause of maternal and fetal mortality and morbidity worldwide. Its pathophysiology remains unclear, but mitochondrial dysfunction and oxidative stress have been implicated. L-Ergothioneine is a naturally occurring, water-soluble betaine, that has demonstrated antioxidant properties. Using the reduced uterine perfusion pressure (RUPP) rat model of pre-eclampsia, this study aimed to define the plasma metabolic profile following treatment with L-Ergothioneine.

Methods

The effect of L-Ergothioneine (ET) treatment was explored using *in vivo* treatment in rats: Sham control (SC, n=5), RUPP control (RC, n=5), Sham +ET (ST, n=5), RUPP +ET (RT, n=5). Differential expression of plasma metabolites were obtained using untargeted liquid chromatography coupled to mass spectrometry. Statistical analysis was performed on normalised data comparing RC to SC, RT to RC, and RT to ST. Metabolites significantly altered (FDR < 0.05) were identified through database search.

Results

We report significantly lower levels of L-palmitoylcarnitine in RC compared to SC, a fatty acyl substrate involved in beta-oxidation in the mitochondria. We report that a metabolite that has been associated with oxidative stress (Glutamylcysteine) was detected at significantly higher levels in RT vs RC and RT vs ST. Five metabolites associated with inflammation were significantly lower in RT vs RC and three metabolites in RT vs ST, demonstrating the anti-inflammatory effects of ET in the RUPP rat model of pre-eclampsia.

Conclusions

L-Ergothioneine may help preserve mitochondrial function by increasing antioxidant levels, and reducing inflammatory responses associated with pre-eclampsia. This study shows the potential of L-Ergothioneine as a treatment for pre-eclampsia.

6.2. Introduction

Pre-eclampsia is a hypertensive disorder of pregnancy, estimated to affect 3-5 % of pregnancies worldwide annually, and leads to over 70,000 maternal deaths and 500,000 fetal and neonatal morbidities every year (Mol et al. 2016; M. A. Brown et al. 2018). The International Society for the Study of Hypertension in Pregnancy (ISSHP) defines pre-eclampsia as *de novo* hypertension after 20 weeks of gestation with or without proteinuria, presence of signs of liver dysfunction, maternal acute kidney injury, hemolysis or thrombocytopenia, neurological features, with or without presence of fetal growth restriction (M. A. Brown et al. 2018). Pre-eclampsia is associated with preterm birth, iatrogenic or spontaneous, and fetal growth restriction (Mol et al. 2016), outcomes which are associated with an increased risk of respiratory and neurological disorders in the fetus. The pathophysiology of pre-eclampsia is still not fully understood, but poor placentation leading to exaggerated oxidative stress and systemic inflammatory response have been implicated (C. W. G. Redman and Sargent 2009; Christopher W. G. Redman and Staff 2015), as have dormant microorganisms (Louise C Kenny and Kell 2018; Kell and Kenny 2016). Deficient placental implantation and failed trophoblastic invasion of the spiral arteries lead to impaired blood perfusion between the myometrium and placenta, in turn leading to placental oxidative stress and an excessive production of reactive oxygen species (ROS) (C. W. G. Redman and Sargent 2009; Christopher W. G. Redman and Staff 2015; Matsubara et al. 2015; Sánchez-Aranguren et al. 2014). ROS are reactive free radicals and other moieties which damage DNA, RNA and proteins; excessive production of ROS can overwhelm natural antioxidant defences and lead to oxidative stress (Birben et al. 2012). Evidence of oxidative stress has been established in the well characterised reduced uterine perfusion pressure (RUPP) rat model of pre-eclampsia (Sedeek et al. 2008; Vaka et al. 2018). Mitochondria are the main cellular source of ROS, and mitochondrial oxidative stress has recently been implicated in the pathophysiology of pre-eclampsia (Vaka et al. 2018; C. McCarthy and Kenny 2016; Williamson et al. 2018). In their study, Vaka *et al.* provided novel evidence that placental ischaemia disrupted mitochondrial function with increased

generation of mitochondrial-specific ROS. Furthermore, they showed that the levels of mitochondrial ROS were elevated in endothelial cells incubated with serum from RUPP rats compared to control rats. In addition, this study showed that endothelial cells incubated with serum from RUPP rats treated with mitochondrial specific antioxidant lead to a decrease in the amout of mitochondrial ROS produced, compared to untreated rats (Vaka et al. 2018).These recent studies indicate that directly targeting mitochondrial dysfunction could be a potential therapeutic treatment for pre-eclampsia.

L-Ergothioneine (ET) is a naturally occurring, water-soluble amino acid and a derivative of histidine, first discovered in the ergot fungus (Borodina et al. 2020; Tanret C 1909). Interestingly, the synthesis of ET has not been observed in animals or higher plants despite evidence of its accumulation in high concentration in animal and human tissues from dietary sources (Cheah and Halliwell 2012); instead it exploits a solute carrier that has been selected for this purpose (Schomig 2005). Antioxidant properties of ET have been demonstrated by the reduction in markers of oxidative stress in blood samples taken from healthy human subjects, who were administered ET (5 or 25mg) once a day for 7 days (Cheah et al. 2017). The radical scavenging properties of ET are established, and this has led scientists to test it as a therapeutic agent in a range of pathologies, such as non-alcoholic fatty liver disease (Cheah et al. 2016). In addition, there is evidence that ET can accumulate in the mitochondrial fraction of hepatic cells in rats (Kawano et al. 1982; E. Smith et al. 2019).

A recent study investigated ET as a therapeutic agent to alleviate clinical symptoms of pre-eclampsia using the well-established reduced uterine perfusion pressure (RUPP) rat model of pre-eclampsia, and determined that its therapeutic properties may in part results from a reduction in mitochondrial-specific ROS (Williamson Rachel et al. 2020). Here, we aimed to explore the plasma metabolic profile of L-Ergothioneine treatment in the RUPP rat model of pre-eclampsia.

6.3. Materials and Methods

6.3.1. Materials and reagents

L-Ergothioneine was provided by Tetrahedron (Paris, France; www.tetrahedron.fr). Liquid chromatography grade methanol and formic acid were purchased from Fisher Scientific (Loughborough, UK). LC-MS glass vials and ultra-performance liquid chromatography (UPLC) column were purchased from Waters (Waters, Wexford, Ireland).

6.3.2. Animals

University College Cork Biological Services Unit supplied and maintained the Sprague Dawley-timed pregnant rats used for the experiment. The animals had free access to food and tap water, and were maintained in a 12-hour light/dark cycle, at 21 ˚C. All the procedures were performed in accordance with National Guidelines and the European Directive 2010/63/EU, under an authorization issued by the Health Products Regulatory Authority Ireland and approved by the Animal Ethics Committee of University College Cork (AE19130/P037).

6.3.3. RUPP procedure

The RUPP procedure has been previously described (Granger et al. 2006; F. P. McCarthy et al. 2011), and is a well-established surgical method used to model the link between hypertension in the pregnant rat and placental ischemia. Briefly, on gestational day (GD) 14, silver clips (0.2 mm ID) were placed on the abdominal aorta (one clip) above the iliac bifurcation under isoflurane anaesthesia; this procedure induced the reduction of blood flow to the uteroplacental unit. An additional two clips (0.1 mm ID) were carefully positioned around the right and left ovarian arteries. As a control, Sham surgery was performed by creating an incision, without placing any clips on either ovarian arteries or abdominal aorta. A chronic indwelling catheter was inserted into the carotid artery on GD18, and mean arterial blood pressure (MABP) was recorded in conscious animals on GD19.

6.3.4. L-Ergothioneine in vivo experimental protocol

The effect of the ET treatment was explored using four experimental groups (**Table 6.1**): Sham control (SC, n=5), RUPP control (RC, n=5), Sham + ET (ST, n=5), RUPP + ET (RT, n=5). From GD11-GD19, ET was added to the drinking water at 25 mg/kg/day, as reported in previous animal studies (Sakrak et al. 2008; Deiana et al. 2004). Blood was collected on GD19 into EDTA vacutainers, these were centrifuged at 2000*g* and 2400*g* for 10 minutes at 4 ˚C, plasma was removed and stored at -80 ˚C for further analysis.

Table 6.1: Description of animal procedures and L-Ergothioneine in vivo experiment.

	Sham Control (n=5)	RUPP Control (n=5)	Sham + ET (n=5)	RUPP + ET (n=5)
Gestational day 11-19	Drinking water	Drinking water	Drinking water + ET	Drinking water + ET
Gestational day 14	Sham surgery	RUPP surgery	Sham surgery	RUPP surgery
Gestational day 19	Caesarean Section, Blood and Tissue Collection Experiment Endpoint			

ET: L-Ergothioneine, administered in drinking water at 25 mg/kg/day.

6.3.5. Plasma sample preparation

Plasma samples were prepared in a randomised order. Samples were taken from -80 ˚C storage and left on ice until thawed. Metabolite extraction was performed based on the protocol by Dunn *et al* (Dunn et al. 2011a), and as follows: cold methanol (300 µl) was added to plasma (100 µl), vortex mixed, and centrifuged at 15,800*g* for 15 minutes at room temperature. Supernatant (200 µl) was transferred in two separate properly labelled Eppendorf tubes, and placed in a centrifugal vacuum evaporator for 4 to 5 hours, without applying any heating. Samples were then reconstituted in water (100µl), vortex mixed and centrifuged at 15,800*g* for 15 minutes at room temperature. Pooled quality control (QC) samples were obtained by taking 20 µl from each sample. Finally, supernatant of all samples and QC were transferred into LC-MS glass vials.

6.3.6. UPLC-MS analysis

Samples were maintained at 4 ˚C during the analysis, analysed in randomised order and in triplicate, on an ultra-high performance liquid chromatography (UPLC) Acquity system coupled to a Synapt G2-S quadrupole time-of-flight (Q-TOF) mass spectrometry system (Waters Corp, Wilmslow, UK). Data were acquired in both positive and negative electrospray ionisation modes (ESI+, ESI-). QC samples were analysed after every 10 samples injections. The chromatographic separation of the

analytes was performed on a BEH C18 column (2.1x100 mm, 1.7 µm) maintained at 50 °C. For both ESI+ and ESI-, the elution gradient used solvent A, a mix of water and formic acid at 0.1 % (v/v), and solvent B, a mix of methanol and formic acid at 0.1 % (v/v). In ESI+, injection volume was 4 µl, and a 22 minutes gradient was applied at 0.40 ml/min, as follows: initial condition at 0 % of B for 1 minute, increased to 100 % of B over 15 minutes, then maintained for 4 minutes, followed by a return to initial conditions of 0 % of B in 2 minutes. In ESI-, injection volume was 7 µl, and a 24 minutes gradient was applied as follows: initial condition at 0 % of B for 2 minutes, increased to 100 % of B over 15 minutes, then maintained for 5 minutes, followed by a return to initial conditions of 0 % of B in 2 minutes.

Data were acquired in "resolution" MSE mode (Bateman et al. 2002; Silva et al. 2005), from 50 to 1,500 Da. Ion data for precursor (low energy) and fragment (high energy) were collected during the same acquisition, with a total cycle time of 0.2 seconds (0.1 second for each). Linear collision energy ramp (20-40 eV) was applied for high energy. Capillary voltage was set to 3.0 kV, sampling cone at 40 V, extraction cone to 5 V. The source temperature was set at 120 °C, and desolvation temperature at 650 °C. Desolvation gas flow rate was set at 800 L/h, and cone gas at 50 L/h. Mass calibration was performed using a sodium formate mix (Waters, Wexford, Ireland), recommended by the manufacturer before each batch analysis. Real time lock mass correction was performed using leucine enkephalin (LeuEnk, 1 ng/µl) mix, injected at 10 µL/min through a lock-spray probe and acquired every 30 seconds.

6.3.7. Data processing

Data processing was performed using Progenesis QI version 2.4 (Nonlinear dynamics, Newcastle, UK), an appropriate pooled QC was selected as the reference to chromatographically align the data, and normalise to all compounds. Data were then peak picked, and considered the adducts corresponding with M+H, M+H-$2H_2O$, M+H-H_2O, M+NH_4, M+Na, M+K, 2M+H (ESI+), or M-H, M-H_2O-H, M-Na-2H, M+K-2H, M+Cl, M+FA-H, 2M-H (ESI-). Downstream statistical analysis was performed using the compound measurements exported from Progenesis QI.

6.3.8. Statistical analysis

Statistical analysis of the data was performed on data normalised to all compounds (using Progenesis QI), and on the averaged intensities of the replicate measures. Data from each group were compared using Student's T test (unpaired, two-tailed, unequal variance) in order to determine which metabolites were significantly altered between the groups. In addition, false discovery rate (FDR) correction using Benjamini- Hochberg (Benjamini and Hochberg 1995), and mean fold change were calculated. Analysis was performed with Excel 2016 Software, and significance was accepted for FDR <0.05. Volcano plots and parallel coordinate plot were created using R statistical software (R Core Team 2013) and the packages ggplot2 (Wickham 2011) and GGally (Schloerke et al. 2011).

6.3.9. Putative annotation of metabolites

The exact mass of significant metabolites (FDR q-value <0.05) were putatively annotated (Salek et al. 2013) by database search using the Progenesis QI identification tool against the Human Metabolome Database (HMDB, version 4) (David S. Wishart et al. 2018). The search parameters were set for an exact mass tolerance 5ppm for the precursor ion, and 10ppm for the fragment ion. Metabolites detected using UPLC-MS are frequently present multiple times due to fragmentation, multiple charging, chemical adduction, or dimerization. Identifications were reported for unique metabolites, after removing duplicates and metabolites of drugs or food, and after checking the retention time. The metabolites reported were grouped into chemical classes using HMDB, and the biological pathways or functions that they are involved with are listed, using HMDB, the small molecule pathway database (SMPDB) (Frolkis et al. 2010), and Kyoto encyclopedia of genes and genomes (KEGG) (KEGG) (**Table 6.2**).

6.4. Results

6.4.1. RUPP control vs Sham control

Using UPLC-Q-TOF-MS, we compared metabolic profiles of RUPP rats (RC) and sham rats (SC) to identify metabolites that were altered in a rat model of pre-eclampsia. The statistical analysis comparing levels of metabolites detected in RC to SC showed that out of 15,024 detected features, 49 metabolites were significant (FDR q-value <0.05); volcano plots are presented in **Figures 6.1A** and **6.1B**. Among these 49 metabolites, 10 were annotated through database search, and belonged in 8 chemical classes (**Table 6.2**): two amino acids, peptides and analogues, one fatty acyl, one glycerolipid, two glycerophospholipids (GPL), two sphingolipids, and two steroids and steroid derivatives. All of these metabolites were detected at lower levels in RUPP control rats compared to sham rats, except for three (two GPL, and one steroid derivative). Interestingly, L-palmitoylcarnitine, a fatty acyl involved in mitochondrial respiration (Schönfeld et al. 2010), was detected at lower levels in RUPP control group with a fold change of -10.35.

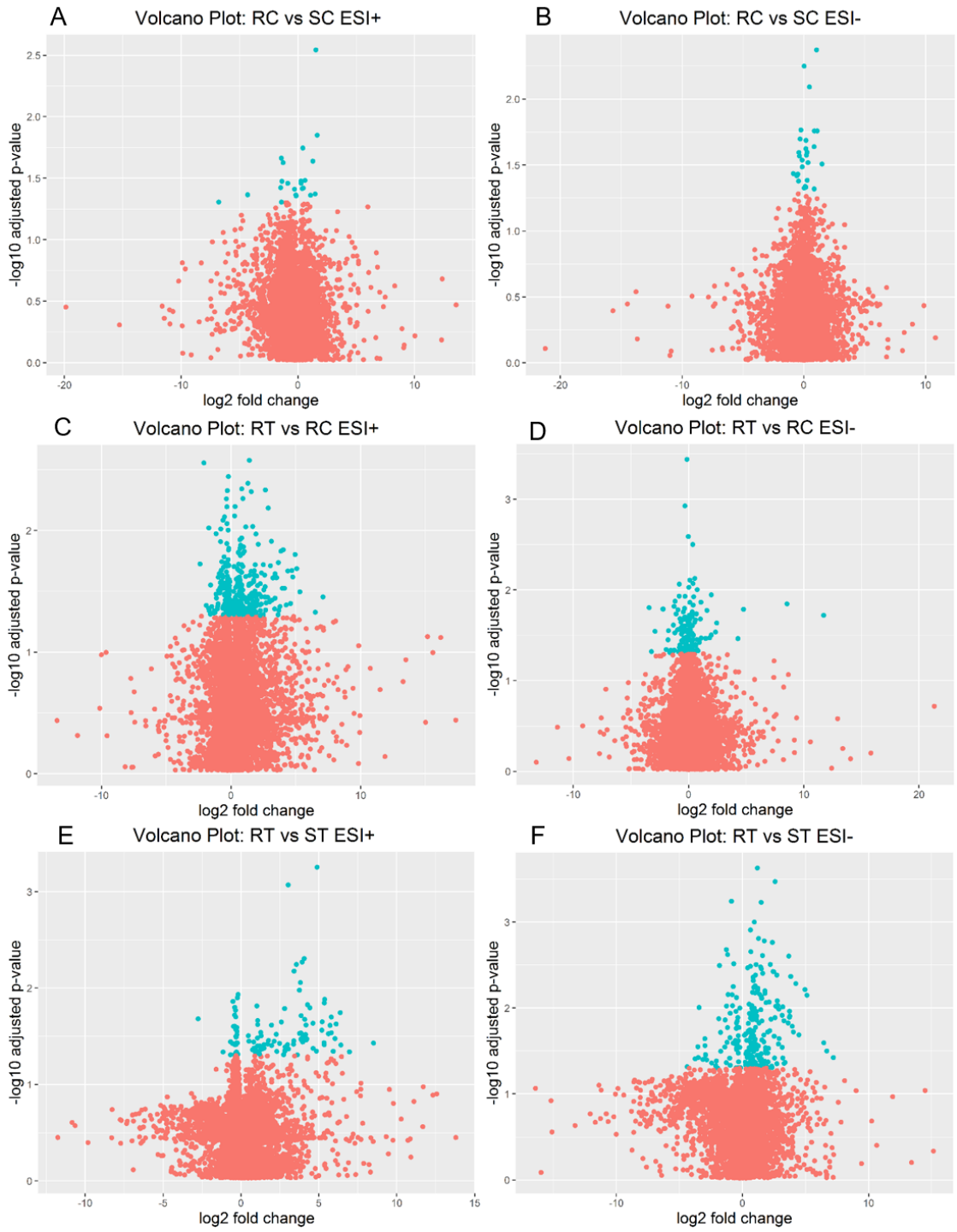


Figure 6.1: Volcano plot of metabolomics data, showing Log2 fold change plotted against Log10 of FDR q-values. (A) RUPP Control (RC) vs Sham Control (SC) in ESI+. (B) RUPP Control vs Sham Control in ESI-. (C) RUPP +ET (RT) vs RUPP Control in ESI+. (D) RUPP +ET vs RUPP Control in ESI-. (E) RUPP+ ET vs Sham+ET (ST) ESI+. (F) RUPP+ET vs Sham+ET ESI-. Features with FDR < 0.05 are plotted in blue.

Table 6.2. Metabolites significantly altered (FDR q-value <0.05) and putatively annotated, for RUPP +ET compared to Sham +ET, RUPP +ET compared to RUPP Control, and RUPP Control compared to Sham Control.

	RUPP +ET vs Sham +ET		RUPP +ET vs RUPP Control		RUPP Control vs Sham Control				
Annotated metabolites	**FDR**	**FC**	**FDR**	**FC**	**FDR**	**FC**	**Chemical class**	**Chemical sub-class**	**Biological function/path way**
9'-Carboxy-alpha-chromanol	1.29E-02	5.95	3.41E-02	2.37			Benzopyrans	1-benzopyrans	vitamin E biosynthesis
Glutamylcysteine OR Cysteinyl-Glutamate	1.37E-02	2.18	8.46E-03	11.11			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Glutathione metabolism

Glycyl-Serine OR Serylglycine	1.53E-02	2.02	3.14E-02	3.57			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Asparaginyl-Lysine OR Lysyl-Asparagine	1.69E-02	-1.28	4.32E-02	-1.21			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
L-Phenylalanine	2.25E-02	6.02	1.95E-02	8.19			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Phenylalanine and Tyrosine Metabolism
L-Glutamic acid	4.72E-02	1.91	4.30E-02	2.64			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Alanine, Histidine, Cysteine Metabolism, Malate-Aspartate Shuttle,

									Ammonia Recycling
Histidinyl-Methionine OR Methionyl-Histidine	4.81E-02	2.27	4.10E-02	4.71			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Asymmetric OR symmetric dimethylarginine	2.29E-02	2.01					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Nonenoylglycine	3.36E-04	5.99					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Hydroxyprolyl-Isoleucine OR Hydroxyprolyl-Leucine	3.12E-03	2.11					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism

Alanyltryptophan OR Tryptophyl-Alanine	1.33E-02	-3.66					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
N,O-Bis-(trimethylsilyl)phenyl alanine	4.88E-02	3.45					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Pyridinoline	4.68E-02	2.04					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Histidinyl-Isoleucine OR Histidinyl-Leucine OR Isoleucyl-Histidine OR Leucyl-Histidine	2.51E-02	-1.39					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
N-Heptanoylglycine OR 3-Hydroxyisoheptanoic	4.54E-02	4.92					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism

acid OR Ethyl 2-hydroxyisovalerate									
Histidinylhydroxyproline	4.87E-02	-1.19					Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
20-COOH-leukotriene E4			2.49E-02	1.56			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Arachidonic acid metabolism
Histidinyl-Tyrosine OR Asparaginyl-Tryptophan OR Tyrosyl-Histidine OR Tryptophyl-Asparagine			2.86E-02	-2.64			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism

Diphthamide			2.80E-02	-2.08			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	
Asparaginyl-Methionine OR Methionyl-Asparagine			4.91E-02	12.39			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Cysteinyl-Glutamate OR Glutamylcysteine			3.73E-02	8.94			Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism
Neuromedin N (1-4)					4.24E-03	-4.56	Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Neuropepetide
Tyrosyl-Phenylalanine OR Phenylalanyl-Tyrosine					4.64E-02	-1.96	Carboxylic acids and derivatives	Amino acids, peptides, and analogues	Protein catabolism

L-Acetylcarnitine			8.22E-03	-1.53			Fatty Acyls	Fatty acid esters	Beta Oxidation of Very Long Chain Fatty Acids, Oxidation of Branched Chain Fatty Acids
L-Palmitoylcarnitine					2.06E-02	-10.35	Fatty Acyls	Fatty acid esters	Fatty acid Metabolism, Fatty acid Degradation
trans-Hexadec-2-enoyl carnitine OR Galactosylsphingosine OR Glucosylsphingosine	4.62E-02	3.32					Fatty Acyls	Fatty acid esters	Sphingolipid metabolism

Azelaic acid	1.69E-02	-2.29					Fatty Acyls	Fatty acids and conjugates	
5-Methyltetrahydrofolic acid (folic acid)	2.37E-02	1.27					Fatty Acyls	Fatty acids and conjugates	
(S)-3-Hydroxytetradecanoyl-CoA	2.73E-02	36.60					Fatty Acyls	Fatty acyl thioesters	Fatty Acid Elongation In Mitochondria, Fatty acid metabolism, Mitochondrial Beta-Oxidation of Long Chain Saturated Fatty Acids
3-Oxotetradecanoyl-CoA	3.89E-02	85.27					Fatty Acyls	Fatty acyl thioesters	Fatty Acid Elongation In

									Mitochondria, Fatty acid metabolism, Mitochondrial Beta-Oxidation of Long Chain Saturated Fatty Acids
S-(2-Methylpropionyl)-dihydrolipoamide-E	2.96E-02	3.66					Fatty Acyls	Fatty amides	Valine, leucine and isoleucine degradation
13-Hydroxyoctadecadienoic acid (13-HODE) OR other fatty acyls - 11 hits	3.93E-02	2.86					Fatty Acyls	Lineolic acids and derivatives	Linoleic acid metabolism

20-Oxo-leukotriene E4	2.35E-02	2.88					Fatty Acyls	Eicosanoids	
Prostaglandin C2 OR 15-deoxy-PGD2 OR other prostaglandins - 8 hits	3.84E-02	2.50					Fatty Acyls	Eicosanoids	Arachidonic Acid Metabolism
20-Trihydroxy-leukotriene-B4 OR Aldosterone OR Cortisone			2.98E-02	6.31			Fatty Acyls OR Steroids and steroid derivatives	Eicosanoids OR Hydroxysteroids	Arachidonic acid metabolism OR Steroid hormone biosynthesis
TG(48:3) OR other TG - 31 hits	3.50E-02	1.48					Glycerolipids	Triradylcglycerols	
DG(36:3) - 2 hits			3.71E-02	1.27			Glycerolipids	Diradylglycerols	Glycerolipid metabolism

DG (35:0) - 3 hits			4.20E-02	1.89			Glycerolipids	Diradylglycerols	Glycerolipid metabolism
DG (39:1) - 3 hits					4.72E-02	-1.48	Glycerolipids	Diradylglycerols	Glycerolipid metabolism
LysoPA - 2 hits	1.38E-02	-2.59					Glycerophospholipids	Glycerophosphates	Glycerolipid metabolism, Glycerophospholipid metabolism
PA - 22 hits			2.89E-02	3.42			Glycerophospholipids	Glycerophosphates	Glycerolipid metabolism, Glycerophospholipid metabolism
PC - 7 hits	3.75E-02	-11.38					Glycerophospholipids	Glycerophosphocholines	Glycerophospholipid metabolism,

									Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC - 8 hits	1.49E-02	-1.39					Glycerophospholipids	Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PC - 9 hits	4.26E-02	1.16					Glycerophospholipids	Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC OR PE - 34 hits	2.86E-02	-1.16					Glycerophospholipids	Glycerophosphocholines OR Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic

									acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC OR PE - 54 hits	2.21E-02	-2.00					Glycerophospholipids	Glycerophosphocholines OR Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PE - 4 hits	4.61E-02	1.22					Glycerophospholipids	Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC OR PE - 18 hits	1.63E-02	15.43					Glycerophospholipids	Glycerophosphocholines OR Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic

									acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC OR PE - 37 hits	6.95E-03	-2.22					Glycerophospholipids	Glycerophosphocholines OR Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PC OR PE - 39 hits	2.96E-02	-1.40					GlycerophospholipidsV	Glycerophosphocholines OR Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PE OR PC - 39 hits			2.07E-02	3.93			Glycerophospholipids	Glycerophosphoethanolamines OR Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic

									acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PE OR PC - 58 hits			4.52E-02	3.53			Glycerophospholipids	Glycerophosphoethanolamines OR Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PE - 3 hits			3.96E-02	2.75			Glycerophospholipids	Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PE OR PC - 16 hits			4.30E-02	-1.42			Glycerophospholipids	Glycerophosphocholines OR Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic

									acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PE OR PC - 36 hits			4.15E-02	-3.61	2.54E-02	9.29	Glycerophospholipids	Glycerophosphoethanolamines OR Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PA - 20 hits			4.54E-02	3.83			Glycerophospholipids	Glycerophosphates	Glycerolipid metabolism, Glycerophospholipid metabolism
PA - 24 hits			3.35E-02	1.43			Glycerophospholipids	Glycerophosphates	Glycerolipid metabolism, Glycerophospholipid metabolism
PS - 10 hits			2.58E-03	1.63			Glycerophospholipids	Glycerophosphoserines	Glycine, serine and threonine metabolism, Glycerophosph

									olipid metabolism
PS - 16 hits			7.44E-03	1.63			Glycerophospholipids	Glycerophosphoserines	Glycine, serine and threonine metabolism, Glycerophospholipid metabolism
PS(36:8)			4.59E-02	1.99			Glycerophospholipids	Glycerophosphoserines	Glycine, serine and threonine metabolism, Glycerophospholipid metabolism
PC - 4 hits			1.96E-02	-1.25			Glycerophospholipids	Glycerophosphocholines	Glycerophospholipid metabolism,

									Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PC - 20 hits			1.88E-02	-5.28			Glycerophospholipids	Glycerophosphocholines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic

									acid metabolism
PE - 7 hits			3.26E-02	-1.82			Glycerophospholipids	Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
PE - 2 hits					2.30E-02	1.02	Glycerophospholipids	Glycerophosphoethanolamines	Glycerophospholipid metabolism, Arachidonic

									acid metabolism, Linoleic acid metabolism, alpha-Linolenic acid metabolism
Hydroxyphenylacetylglycine OR Hypoxanthine			4.12E-02	2.73			Imidazopyrimidines OR Carboxylic acids and derivatives	Purines and purine derivatives OR Amino acids, peptides, and analogues	Tyrosine metabolism OR Purine Metabolism
Decanoylcholine	2.00E-02	1.20					Organonitrogen compounds	Quaternary ammonium salts	
N-Acetyl-D-glucosaminyldiphosphodolichol	4.95E-02	4.80					Organooxygen compounds	Carbohydrates and carbohydrate conjugates	

3-Sialyl-N-acetyllactosamine OR 6-Sialyl-N-acetyllactosamine	3.03E-02	17.71					Organooxygen compounds	Carbohydrates and carbohydrate conjugates	
Quinone			1.19E-02	-1.28			Organooxygen compounds	Carbonyl compounds	Riboflavin Metabolism, Pyrimidine Metabolism
Spermidine monoaldehyde 1			4.52E-02	1.67			Organooxygen compounds	Carbonyl compounds	Arginine and proline metabolism, beta-Alanine metabolism, Glutathione metabolism, Bile secretion

N-Acetylneuraminic acid			3.43E-02	3.02			Organooxygen compounds	Carbohydrates and carbohydrate conjugates	Amino Sugar and Nucleotide Sugar Metabolism
Hydroxyphenyllactic acid	1.37E-02	2.18					Phenylpropanoic acids		
3-Hydroxy-4-methoxyphenyllactic acid OR other polyphenol metabolites - 65 hits	9.39E-03	-1.73					Phenylpropanoic acids		
Formyl-CoA			4.64E-03	6.27			Purine nucleotides	Purine ribonucleotides	Oxidation of Branched Chain Fatty Acids, Glyoxylate and dicarboxylate metabolism

Ceramide (42:0)	2.27E-02	7.47					Sphingolipids	Ceramides	Sphingolipid metabolism
Ceramide (44:1)	1.49E-02	2.50					Sphingolipids	Ceramides	Sphingolipid metabolism
Ganglioside GA2 (30:0) OR Trihexosylceramide (30:1)	2.24E-02	37.00					Sphingolipids	Glycosphingolipids	Highly important in immunology
SM(32:1)	3.41E-02	1.20					Sphingolipids	Phosphosphingolipids	Sphingolipid metabolism, Sphingolipid signaling pathway, Necroptosis
Glucosylceramide (42:1)			4.08E-02	1.33	3.70E-02	-1.33	Sphingolipids	Glycosphingolipids	Sphingolipid metabolism

Ganglioside GM3 (32:0)			4.19E-02	-1.67			Sphingolipids	Glycosphingolipids	Highly important in immunology
Galabiosylceramide (44:1) OR Lactosylceramide (44:1)			3.12E-03	2.20			Sphingolipids	Glycosphingolipids	Sphingolipid metabolism
Glucosylceramide (36:1) OR Galactosylceramide (36:1)			1.98E-02	1.17			Sphingolipids	Glycosphingolipids	Sphingolipid metabolism
Ganglioside GM3 (40:2)			3.82E-02	3.23			Sphingolipids	Glycosphingolipids	Highly important in immunology
SM(44:1)					3.88E-02	1.16	Sphingolipids	Phosphosphingolipids	Sphingolipid metabolism, Sphingolipid

									signaling pathway, Necroptosis
3alpha,21-Dihydroxy-5beta-pregnane-11,20-dione	2.76E-02	2.41	3.71E-02	1.88	2.00E-02	-1.36	Steroids and steroid derivatives	Hydroxysteroids	Steroid hormone biosynthesis
17alpha,21-Dihydroxy-5beta-pregnane-3,11,20-trione OR Cortisol OR 18-Hydroxycorticosterone	3.06E-02	1.53					Steroids and steroid derivatives	Hydroxysteroids	Steroid hormone biosynthesis
Chenodeoxycholic acid sulfate OR Chenodeoxycholic acid 3-sulfate OR	6.18E-03	2.29					Steroids and steroid derivatives	Bile acids, alcohols and derivatives	Fat digestion and absorption, Vitamin

Ursodeoxycholic acid 3-sulfate									digestion and absorption
Glycoursodeoxycholic acid OR Deoxycholic acid glycine conjugate OR Chenodeoxycholic acid glycine conjugate OR Chenodeoxyglycocholic acid	3.48E-02	-7.61					Steroids and steroid derivatives	Bile acids, alcohols and derivatives	Fat digestion and absorption, Vitamin digestion and absorption
Estrone glucuronide	3.97E-02	2.09					Steroids and steroid derivatives	Steroidal glycosides	Steroid hormone biosynthesis
25-Hydroxyvitamin D3-26,23-lactone			4.77E-02	2.09			Steroids and steroid derivatives	Vitamin D and derivatives	Steroid hormone biosynthesis

Dehydroepiandrosterone 3-glucuronide OR Dehydroisoandrosterone 3-glucuronide			3.60E-02	3.54			Steroids and steroid derivatives	Steroidal glycosides	Steroid hormone biosynthesis
2-hydroxyethinylestradiol			2.51E-02	1.34			Steroids and steroid derivatives	Estrane steroids	Steroid hormone biosynthesis
5beta-Pregnanediol					4.32E-02	1.03	Steroids and steroid derivatives	Pregnane steroids	Steroid hormone biosynthesis
trans-Resveratrol 4'-O-glucuronide OR cis-Resveratrol 4'-O-glucuronide			1.17E-02	76.61			Stilbenes	Stilbene glycosides	Stilbenoid, diarylheptanoid and gingerol biosynthesis

FDR: false discovery rate q-value; FC: fold change.

6.4.2. RUPP +ET vs RUPP control

To identify metabolites that were altered following L-Ergothioneine treatment in the RUPP rat model of pre-eclampsia, we compared metabolic profiles of RUPP rats treated with ET (RT) to profiles of RUPP rats not treated (RC). Following statistical analysis, and out of 15,034 detected features, we reported 529 metabolites as significantly altered (FDR q-value <0.05), volcano plots are presented in **Figures 6.1C and 6.1D**, and 45 metabolites were annotated through database search (**Table 6.2**). The metabolites annotated belonged to the following chemical classes: fourteen glycerophospholipids (GPL), eleven amino acids, peptides, and analogues, five sphingolipids, four steroid and steroid derivatives, two fatty acyls, two glycerolipids, one benzopyran, one imidazopyrimidine, three organooxygen compounds, one purine nucleotide, and one stilbene. The majority of these metabolites were detected at higher levels in the RUPP+ET group compared to RUPP control group. Out of the fourteen GPL detected, eight were phosphatidylcholines (PC) or phosphatidylethanolamines (PE), and are involved in arachidonic acid metabolism, linoleic acid metabolism, and alpha-linolenic acid metabolism. In addition, two metabolites involved in arachidonic acid metabolism were differentially detected between RT and RC: 20-Trihydroxy-leukotriene-B4, and 20-COOH-leukotriene E4. The second chemical class that showed the most altered metabolites was amino acids, peptides, and analogues, which are involved with protein catabolism, and two of these identified amino acids are involved with glutathione metabolism (Glutamylcysteine or Cysteinyl-Glutamate), and one with the malate-aspartate shuttle (L-Glutamic acid). In addition, two gangliosides were detected (Ganglioside GM3 (32:0) with a fold change of -1.67, and Ganglioside GM3 (40:2) with a fold change of 3.23).

6.4.3. RUPP+ET vs Sham+ET

Similarly, we compared metabolic profiles of RUPP+ET to profiles of Sham+ET, and following statistical analysis of the 15,024 detected features, we reported 368 significant metabolites (FDR q-value <0.05), volcano plots are presented in **Figures**

6.1E and 6.1F. 50 were annotated following database search (**Table 6.2**). The annotated metabolites belong in 10 chemical classes: fifteen amino acids, peptides, and analogues, ten glycerophospholipids (GPL), nine fatty acyls, five steroids and steroid derivatives, four sphingolipids, two organooxygen compounds, two phenylpropanoic acids, one benzopyran, one organonitrogen compound, and one glycerolipid. The majority of these metabolites (36) were detected at significantly higher levels in RUPP+ET group compared to Sham+ET group. Similar to the comparison of RUPP+ET to RC, among the 10 GPL detected as significanlty altered in RUPP+ET compared to Sham+ET, nine are PC or PE.

In addition, 20-Oxo-leukotriene E4, a metabolite of Leukotriene E4 (LTE4) following lipid oxidation, was detected at higher levels in RUPP+ET rats. Furthermore, two fatty acyls showed very high fold change in RUPP+ET compared to Sham +ET, (S)-3-Hydroxytetradecanoyl-CoA (FC 36.6) and 3-Oxotetradecanoyl-CoA (FC 85.27). Both of these fatty acids are involved with fatty acid metabolism, mitochondrial beta-oxidation of long chain saturated fatty acids, and fatty acid elongation in the mitochondria. In addition, Ganglioside GA2 (30:0) (or Trihexosylceramide (30:1)) was detected with a fold change of 37 in RUPP+ET compared to Sham+ET. Five of the amino acids and peptides detected as significantly altered between RUPP+ET and RC were also significantly altered in RUPP+ET compared to Sham+ET, most of which were detected at higher levels in RUPP+ET compared to Sham+ET, including L-Glutamic acid (FC 1.91) and Glutamylcysteine or Cysteinyl-Glutamate (FC 2.18).

Three metabolites were detected as significantly altered both in RUPP control vs Sham and in RUPP+ET vs RUPP control, with opposite fold changes: one GPL (a phosphatidylethanolamine or phosphatidylcholine), one sphingolipid (glucosylceramide (42:1)). One steroid derivative (3alpha,21-Dihydroxy-5beta-pregnane-11,20-dione) was detected as significantly altered in all comparisions: RUPP+ET compared to Sham+ET (FDR=2.76E-02, FC 2.41), RUPP+ET compared to RUPP Control (FDR=3.71E-02, FC 1.88), and RUPP Control compared to Sham Control (FDR=2.00E-02, CF -1.36). A parallel coordinate plot showing the levels of 10 metabolites of interest in the fours experimental groups is shown in **Figure 6.2**, and

demonstrates that for most of these metabolites, the levels in the RUPP+ET group are similar to those in the Sham Control group.

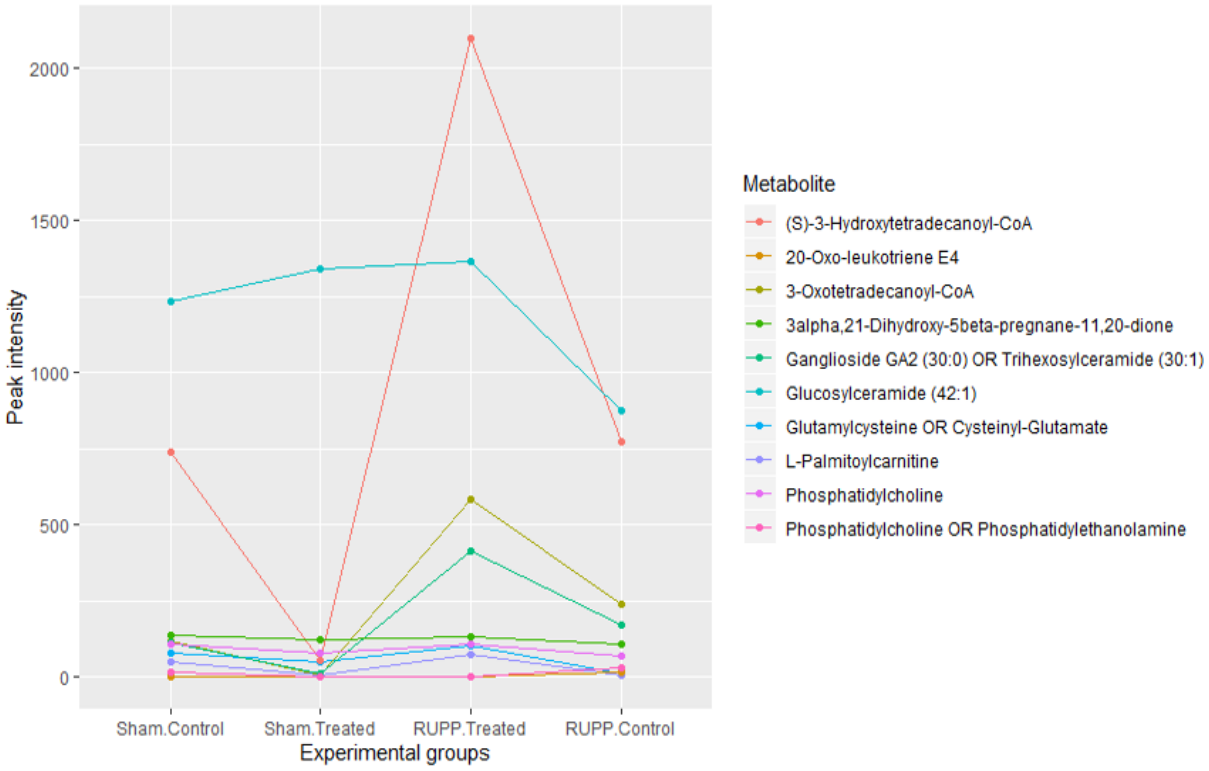


Figure 6.2: Parallel coordinate plot showing the averaged and normalised peak intensity of ten metabolites of interest in Sham Control, Sham +ET, RUPP Control and RUPP +ET rats.

In addition, L-Ergothioneine and two of its metabolites (Hercynine, and S-methyl-L-Ergothioneine) were detected at similar relative levels in the plasma of all groups of animals in ESI+ (**Figures 6.3** and **6.4**). This presumably reflects the fact that it is excreted only very slowly so accumulates in the tissues.

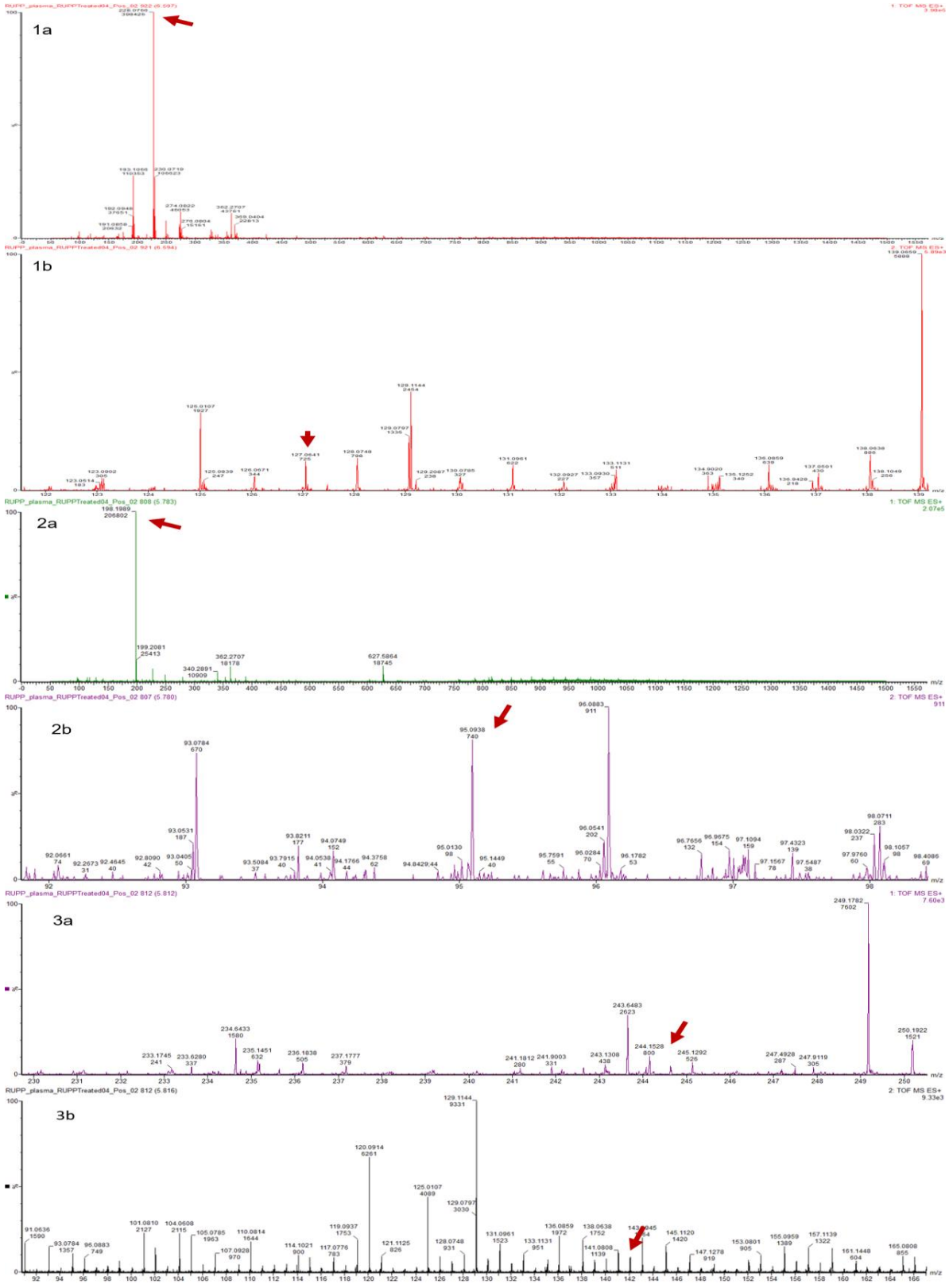


Figure 6.3: Mass spectra of L-Ergothioneine (1a, 1b), L-Hercynine (2a, 2b), and S-methyL-Ergothioneine (3a, 3b), detected in ESI+ mode using UPLC-Q-TOF-MS analysis. Spectra show m/z and intensity of peaks detected. 1a: MS1 scan showing

parent ion of L-Ergothioneine at 228.0766 m/z, and 1b: MS2 scan showing its product ion at 127.0641 m/z; 2a: MS1 scan showing parent ion of L-Hercynine at 198.1989 m/z, and 2b: MS2 scan showing its product ion at 95.0938 m/z; 3a: MS1 scan showing parent ion of S-methyL-Ergothioneine at 244.1528 m/z, and 3b: MS2 scan showing its product ion at 141.0808m/z.

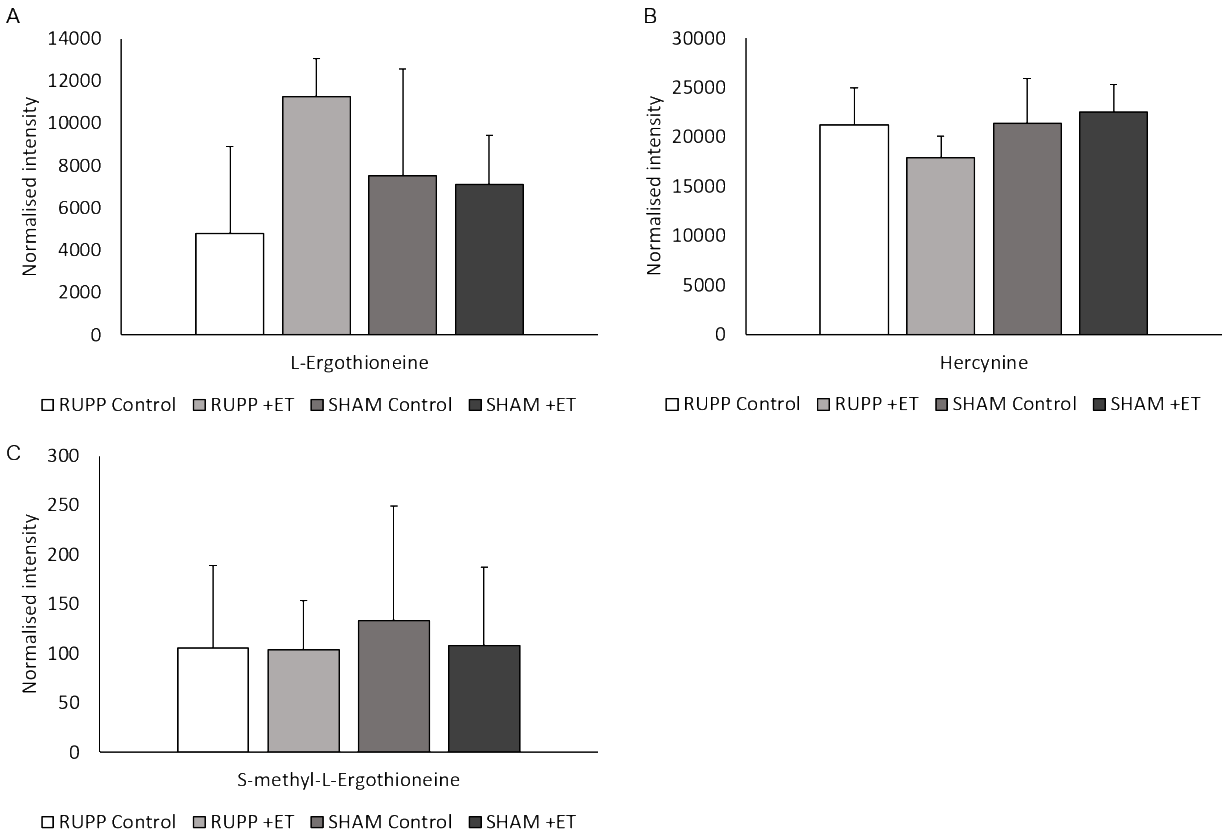


Figure 6.4: Plasma levels of L-Ergothioneine (A), Hercynine (B), and S-methyl- L-Ergothioneine (C). Levels are shown as normalised intensity as detected in ESI+ UPLC-Q-TOF-MS in plasma samples of RUPP Control, RUUP +ET, Sham Control and Sham +ET groups. No significant differences in levels observed between the groups. All values are expressed with standard error.

6.5. Discussion

The present study showed evidence of mitochondrial dysfunction in the plasma metabolite profile in the RUPP rat model of pre-eclampsia. This was evident from the lower levels of L-palmitoylcarnitine in RUPP+ET compared to RUPP Control (FC -10.35), which is associated with beta-oxidation in the mitochondria; from the higher levels of 3-Hydroxytetradecanoyl-CoA (FC 36.6) and 3-Oxotetradecanoyl-CoA (FC 85.27) in RUPP+ET compared to Sham+ET; and from the detection of several significantly altered sphingolipids, which are involved in various cellular processes, such as apoptosis and inflammation (Baker et al. 2016; Funke et al. 2013; Underwood et al. 2006).

This study showed that L- palmitoylcarnitine was detected at significantly lower levels in RUPP Control compared to Sham Control rats (FC -10.35). Under physiological conditions, the mitochondria uses oxidative phosphorylation and the electron transport chain to produce ATP. However, it has been shown that pre-eclampsia is associated with mitochondrial dysfunction and oxidative stress (Williamson et al. 2017; Williamson et al. 2018), and under such conditions, it is hypothesied that L- palmitoylcarnitine could be used as a mediator of mitochondrial respiration which could explain its lower levels in the RUPP Control rats, a model for pre-eclampsia, compared to healthy Sham Control rats. It has been shown that RUPP rats have increased evidence of mitochondrial dysfunction compared to sham controls (Vaka et al. 2018; Williamson Rachel et al. 2020). Williamson *et al.* have recently shown *in vivo* that mitochondrial-specific production of H_2O_2 in kidney tissue is increased in RUPP rats compared to sham control rats (Halliwell et al. 2018), which further demonstrated that mitochondrial dysfunction is associated with the pathophysiology of pre-eclampsia. This study also showed that the RUPP rats treated with ET (RUPP+ET) presented altered levels of fourteen glycerophospholipids (GPL) compared to RUPP rats (RC). Eight of these GPL are part of the arachidonic acid, linoleic acid, and alpha-linolenic acid metabolism pathways, which are used as precursors in the biosynthesis of several metabolites, such as eicosanoids or prostaglandins, involved in multiple cellular processes, including inflammation (Baker et al. 2016). Here, we reported that five of these metabolites were detected at lower

levels in RT rats compared to RC rats, demonstrating the anti-inflammatory effects of ET in the RUPP rat model of pre-eclampsia. In addition, L-glutamic acid was detected at higher levels in RT rats compared to RC. This amino acid is involved in the malate-aspartate shuttle system, or the malate shuttle, which is an essential system used by functional mitochondria allowing electrons to effectively move across the impermeable membrane between the cytoplasm and the mitochondrial matrix. Another metabolite of interest detected at higher levels in RUPP+ET rats is glutamylcysteine. This dipeptide is associated with glutathione metabolism, and a precursor of glutathione (GSH), an antioxidant and free radical scavenger (Guoyao Wu et al. 2004). Taken together, higher levels of glutamylcysteine in RUPP+ET rats indicates that ET can elevate antioxidant activity in the RUPP model of pre-eclampsia. One detected metabolite was annotated for glutathione, however it was not dected as significantly altered between the four experimental groups. We can hypothesise that L-Ergothioneine treatment in the RUPP rats helps alleviate oxidative stress by using a pathway alternative to glutathione, which is in line with the discovery of a solute carrier dedicated to the transport of L-Ergothioneine, solute carrier family 22 member 4 (SLC22A4) (Schomig 2005).

Furthermore, this study showed that two fatty acyls were detected as significantly altered and with very high fold change in RUPP+ET compared to Sham+ET, (S)-3-Hydroxytetradecanoyl-CoA (FC 36.6) and 3-Oxotetradecanoyl-CoA (FC 85.27). Both of these fatty acyls are involved with fatty acid metabolism, mitochondrial beta-oxidation of long chain saturated fatty acids, and fatty acid elongation in the mitochondria, and further the hypothesis of mitochondrial dysfunction during pre-eclampsia (Vaka et al. 2018; C. McCarthy and Kenny 2016; Williamson et al. 2018). Two metabolites involved in arachidonic acid or linoleic acid metabolism pathways were significantly elevated in RUPP+ET compared to Sham+ET: Prostaglandin C2, 13-Hydroxyoctadecadienoic acid (13-HODE). In addition, Ganglioside GA2 (30:0) (or Trihexosylceramide (30:1)), was detected with a fold change of 37 in RUPP+ET compared to Sham+ET. Taken together, the significant alterations observed when comparing RUPP+ET to RUPP Control, and RUPP+ET to Sham Control tend to confirm the effects of L-Ergothioneine on oxidative stress and inflammatory response.

The low number of samples per rat groups (n=5) is a limitation of the present study. However, the experiments were designed to obtain good quality data with the use of quality control samples and analysing the samples in a randomised order, as is recommended by metabolomics guidelines (Broadhurst et al. 2018). In addition, this study showed that metabolic profiling of plasma from RUPP rat model of pre-eclampsia was able to detect pathways altered by L-Ergothioneine treatment.

A previous metabolomics study aiming to identify early biomarkers of prediction of small for gestational age was performed on plasma samples taken from RUPP rats, human cord blood, and human maternal plasma sample collected at 15 weeks of gestation (Horgan et al. 2011). In this study, Horgan *et al.* aimed to validate in RUPP samples the panel of potential metabolite biomarker determined in human samples. Doing so, they confirmed eleven out of nineteen metabolites were also significantly altered in RUPP plasma samples. This panel includes an amino acid, a carnitine or vitamin D derivative, a fatty acid, a glycerolipid, and four glycerophospholipids. The majority of these metabolites were detected at higher levels in RUPP samples than in control rat samples. In our study, we observed that similar metabolite classes were significantly altered when comparing RUPP Control to Sham Control, and RUPP+ET to Sham+ET profiles. However, these metabolites show opposite trends in our study compared to what Horgan *et al.* reported. This said, these two studies were designed with major differences. Horgan *et al.* focused on small for gestational age and only reported metabolites in RUPP plasma samples, while in our study we focused on the pre-eclampsia phenotype and reported the metabolites significantly altered between RUPP control and Sham control, as well as after treatment with L-Ergothioneine by comparing profiles of RUPP+ET to RUPP control, and RUPP+ET to Sham+ET.

The strength of our data compared to previous studies investigating the pharmacological properties of L-Ergothioneine is that we examined its effects in the RUPP rat model of pre-eclampsia, and not only in healthy subjects as previously reported (Cheah et al. 2017; Forster et al. 2015). The data presented in this study correlate with data previously reported by Cheah *et al.* (Cheah et al. 2017) and Tang

et al. (Tang et al. 2018), where both precursor ion and daughter fragment ions of L-Ergothioneine, Hercynine, and S-methyl-L-Ergothioneine were detected; however, there were no significant differences between the plasma levels in all groups in the present study. As previously reported, L-Ergothioneine accumulates in various organs and tissues in animals (Cheah and Halliwell 2012), including humans (Cheah et al. 2017), especially in the liver, as well as in kidney, heart, spleen, intestines, eye, and brain tissues (Tang et al. 2018). It has been reported that both brain tissue and urine from newborns contain L-Ergothioneine (Graham et al. 2017; Vavich and Howell 1970), which implies this molecule is able to cross the blood-placental barrier (Halliwell et al. 2018). A neuroprotective role for L-Ergothioneine has also been suggested by a number of *in vitro* and *in vivo* studies (Halliwell et al. 2018) however, more studies are needed to determine the effects on the newborns. Administration of pure L-Ergothioneine to animals (Forster et al. 2015) or healthy human subjects (Cheah et al. 2017) has shown no adverse effects, and there is thus great potential to use this naturally occurring amino acid as a potential treatment for pre-eclampsia in a clinical trial.

In recent years, metabolomics techniques have become crucial tools in drug discovery (David S Wishart 2016). A number of recent metabolomic alterations detected in pathologies, such as Alzheimer's diseases (Cai et al. 2012), frailty (Picca et al. 2019), atherosclerosis (J. M. Brown and Hazen 2015), cardiovascular disease (E. Smith et al. 2019), or autism and schizophrenia (Shaw 2010), have led to new hypotheses on their aetiologies and new drug targets. The study of pathologies using metabolomics approaches allows for the detection of elevated or lower levels of specific metabolites, which are regulated by enzymes for either biosynthesis or used as a substrate in a biochemical reaction, and these enzymes can be targeted and new drug treatments can be tested (David S Wishart 2016).

In conclusion, L-Ergothioneine may mediate its therapeutics properties in part by preserving mitochondrial function, increasing antioxidant levels in addition to dampening the inflammatory response evident in pre-eclampsia. Furthermore, this study highlights the potential of L-Ergothioneine as a possible treatment for pre-eclampsia.

Chapter 7:

General Discussion

7.1. Summary of main findings

This thesis was conducted with the overarching aim of applying metabolomics techniques to study the pathophysiology of adverse pregnancy outcomes.

Overall, I found metabolomics to be a promising tool to investigate at risk pregnancy states. Notably, this thesis showed that altered glycerophospholipid levels in maternal plasma samples are associated with small for gestational age in babies (SGA), and with spontaneous preterm birth (sPTB). This thesis also investigated and validated potential lipid biomarkers predictive of sPTB. Lastly, this thesis studied the metabolomic effects of a potential therapeutic agent for the treatment of pre-eclampsia in a rat model of pre-eclampsia, and showed that metabolomics techniques can be used to help understand potential therapeutic agents.

Chapter 2 presents a systematic review examining the predictive accuracy of metabolomics for SGA babies. My findings showed that to date and despite research efforts, no robust predictive biomarker of SGA has been discovered through the use of metabolomics techniques. However, my systematic study highlighted the clear alteration of lipid metabolism in this adverse pregnancy outcome, which warrants further investigation (Leite et al. 2019). Indeed, several fatty acids (omega-3, omega-6 (van Eijsden et al. 2008a)) and glycerophospholipids, (lysophosphatidylcholine, sphinganine 1-phosphate, or sphingosine 1-phosphate (Horgan et al. 2011)), were found to be altered in plasma samples of women with SGA pregnancies as compared to women who had uncomplicated pregnancies. However, none of these potential biomarkers have been validated to date. For this reason, I chose to focus on the study of lipid metabolites in plasma samples in **Chapter 3**.

In **Chapter 3**, I investigated the pathophysiology of SGA (customised birthweight <10th centile), using urine and plasma samples taken at 20 weeks of gestation, from women participating in the SCOPE birth cohorts in Cork and Auckland. My findings from the lipidomics experiment showed that women who delivered an SGA baby had higher plasma levels of glycerophospholipids compared to their matched controls in the Cork population. Moreover, the metabolomics experiments performed on urine samples showed lower levels of metabolites associated with detoxification, and

nutrient transport and absorption, in women with SGA pregnancies compared to their matched controls in both Cork and Auckland populations. The novelty of this study was the analysis of both plasma and urine samples from the same participants. Taken together, urine and plasma analyses results were consistent with previous literature and confirm that placental dysfunction is involved in the pathophysiology of SGA. These findings further aid understanding of the pathophysiology of SGA by confirming that higher levels of glycerophospholipds are involved.

In **Chapter 4**, lipidomics discovery analysis was used to study the pathophysiology of sPTB at 20 weeks of gestation in the Cork population of the SCOPE birth cohort. This study revealed for the first time that reduced levels of plasma glycerophospholipids and sphingolipids at 20 weeks of gestation are associated with sPTB (<34 weeks) (Morillon et al. 2020b), and provided further insight into the aetiology of sPTB. Our findings also suggest that lipids could be potential biomarkers of this adverse pregnancy outcome.

In **Chapter 5**, longitudinal biomarker discovery and targeted validation analyses were performed to identify plasma lipid biomarkers predictive of sPTB in women who participated in the SCOPE birth cohort in Cork and Auckland. Two of our lipid targets were detected with altered plasma levels at 15 weeks of gestation (phosphatidylcholine (15:0/22:6)), and at 20 weeks of gestation (triacylglycerol (18:3/18:2/18:3)). These two lipids did not accurately predict sPTB; however, the evidence of altered levels of phospholipids associated with sPTB gives further insights into the pathophysiology of this complex and multifactorial pregnancy complication.

Finally, in **Chapter 6**, the effects of L-Ergothioneine, a known antioxidant and potential therapeutic treatment for pre-eclampsia, on the plasma metabolite profile of the reduced uterine perfusion pressure (RUPP) rat model of pre-eclampsia was investigated. Using untargeted metabolomics analysis, this chapter provided evidence that L-Ergothioneine may help preserve mitochondrial function by increasing antioxidant levels, and reducing inflammatory responses associated with pre-eclampsia. Such results demonstrated that metabolomics techniques can be used to investigate potential therapeutic agents and their effects on the metabolome.

7.2. Clinical implications of metabolomics findings

This thesis offers evidence that applying metabolomics techniques to study the pathophysiology of adverse pregnancy outcomes can generate valuable insight into their pathophysiology and can suggest potential predictive biomarkers, as well as demonstrating the translational potential of clinical metabolomics. Indeed, by using metabolomics techniques it is possible to determine subtle alterations at the level of the metabolome. This allows for the study of disease phenotypes, at specific time points, in minimally invasive way (i.e. urine, plasma, serum samples). These factors make metabolomics well suited for clinical research and the longitudinal monitoring of specific phenotypes (Souza et al. 2019).

The precise aetiology leading to the onset of SGA and sPTB, two complex adverse pregnancy outcomes, are not fully understood. The body of work presented in this thesis contributed to further understanding their pathophysiology, and offered further evidence that the metabolome of women whose pregnancies were subsequently complicated by SGA or sPTB was altered compared to that of women who had uncomplicated pregnancies. This thesis showed that glycerophospholipids may be involved in the pathophysiology of these two pregnancy complications. Altered levels of circulating glycerophospholipids in maternal plasma samples at 15 or 20 weeks of gestation could reflect its impaired placental transport, which may lead to sPTB or SGA. However, more experiments are needed to understand the specific mechanisms involved in the altered lipid levels measured in maternal circulation.

Current diagnosis of SGA and sPTB remains difficult as no predictive biomarker has been found robust enough for routine clinical use. This thesis has yielded promising results that suggested lipids could be used as potential biomarkers. Future research could build on these observations by undertaking validation, paving the way for the development of new diagnostic tools and screening tests for SGA and sPTB. Such an example of metabolic biomarker used as a screening tool is PrePsia^{TM}, a commercial screening test for pre-eclampsia produced by Metabolomics Diagnostics

(www.metabolomicdiagnostics.com), and currently the subject of clinical trials. This test screens for elevated plasma levels of marinobufagenin (MBG), a cardiotonic steroid, using UPLC/MSMS method (Lenaerts et al. 2018).

Furthermore, the evidence presented in this thesis, showed that metabolomics is a powerful tool to investigate the effects of potential drug treatments on the metabolome.

7.3. Thesis strengths

The main strength of this thesis is the use of samples taken from the well-defined Screening for Pregnancy Endpoint (SCOPE) cohort study, and the uniqueness of the populations studied (see **sections 1.4.5.1 and 1.4.5.2**). Using samples from the SCOPE study for the work presented here, allowed to analyse samples taken from very distinct populations, which makes the findings presented here stronger.

In order to maintain the highest data and sample collection quality in the SCOPE study, all participating centres followed the same standard operating procedures, which included sample collection and storage. The prospective collection of data allowed us to rigorously match our case samples to controls, and ensured we analysed high quality samples with a robust experimental design (North et al. 2011). Indeed, this rigorous matching of our samples ensures that our findings are representative of the adverse pregnancy outcome studied, and not that of a sub-phenotype where several pregnancy complications are observed (e.g. pre-eclampsia and SGA, or SGA and sPTB).

Another strength of this thesis is that through several experimental designs, we have shown how clinical metabolomics benefitted our understanding of SGA, sPTB, and pre-eclampsia. Indeed, the application of metabolomics techniques can be applied using various experimental designs, depending on which is the most appropriate to test the hypothesis (Dunn et al. 2011b).

Clinical metabolomics biomarker discovery requires interdisciplinary knowledge, and requires a range of expertise including bio banking, clinical input, epidemiology,

industry and state of the art instrumentation, and bioinformatics (M. Brown et al. 2005). The work presented in this thesis was performed using robust and optimised preparation and analysis methods, which we developed in partnership with our industry partner, Waters Corporation. Moreover, through this partnership, we were able to perform the analyses on state of the art mass spectrometers (Waters Synapt G2-S Q-TOF-MS; Waters Xevo TQ-MS), and bioinformatics software (Progenesis QI) for the data processing and metabolites identifications. In addition, the statistical analyses of the experiments described in this thesis, were performed with the collaboration of a consultant statistician.

Lastly, validation analyses were undertaken to try and confirm the findings from our discovery analysis (**Chapter 5**). Indeed, longitudinal validation experiment in demographically distinct population was performed using samples from Cork and Auckland, collected at 15 weeks of gestation, 20 weeks of gestation, and at time of delivery. To our knowledge, this is the first longitudinal validation study for biomarkers of sPTB. This enables the exploration and monitoring of the levels of potential predictive biomarkers over time and could allow for a better path of care for patients in a clinical setting.

7.4. Thesis limitations

The main limitation of this thesis was the sample size and power of the studies performed, despite having access to SCOPE. However, the experiments presented in this thesis were designed to offer sufficient statistical power (**section 1.4.5.2**). SGA babies are estimated to occur in 10 % of pregnancies (D. Sharma et al. 2016b); preterm birth (PTB) is estimated to affect 5-9 % of pregnancies in Europe, with approximately 31-50 % of all PTB being spontaneous (sPTB) (Slattery and Morrison 2002), which is in agreement with the number of cases observed in SCOPE Cork and Auckland cohorts. Our nested case-control experiments were designed to be robust enough, as we applied robust statistical analysis to help address the sample size. The statistical analysis included empirical Bayesian methods (Smyth 2004) with stringent false discovery rate cut-off (5 %). We estimated that given: the univariate power

calculations (based on similar clinical studies (Auray-Blais et al. 2011; Menon et al. 2014)), the strict collection/processing protocols, the close matching of cases and controls, and the relatively homogeneous target population, then for a metabolomics discovery study the number of samples collected was sufficient for the aims of this study. In addition, several discovery articles have been published with smaller sample size than ours, with 15 samples (Auray-Blais et al. 2011), or 50 samples analysed (Menon et al. 2014). Furthermore, the findings obtained at 15 weeks of gestation were reproduced in samples taken at 20 weeks of gestation (**Chapter 5**), and are in agreement with the literature (E. C. Considine et al. 2019), which demonstrates that the detected altered lipid levels were unlikely due to false discovery.

The second limitation of this thesis was metabolite identifications, which is a major bottle neck in the field of metabolomics (Dunn et al. 2013). Using high specificity platforms, such as the mass spectrometer used for the untargeted analyses presented in this thesis, several thousands of metabolites were detected in each analysis (Dunn et al. 2011b). Obtaining the absolute identification of all detected metabolites of interest (e.g. detected in statistically different levels in the groups compared) would be extremely time consuming and onerous (e.g. cost for custom made standards). This is especially true when most metabolites are unknown and would necessitate *de novo* identification. However, the use of curated metabolites libraries and databases for *in silico* identification is helpful, and it is the most common method to identify metabolites in untargeted discovery analyses to date (M. Brown et al. 2009; Draper et al. 2009). Following metabolomics guidelines (Salek et al. 2013), the latter method was used in the analyses presented in this thesis to obtain putative annotation of metabolites.

The third limitation of this thesis was the inherent technical and biological variability of the experiments presented. This variability has multiple origins, however, the study designs, sample preparations and instruments optimisations all contributed to minimise them to be negligible compared to the biological variance of the samples analysed (see **Chapter 1**, **sections 1.4.5.1** and **1.4.5.3**) (Broadhurst et al. 2018). Untargeted assays yield thousands of detected metabolites with unknown chemical identity, which makes the use of internal standards for all detected metabolites, and

the use of calibration curves impossible, contrary to targeted analyses (Broadhurst et al. 2018). The experiments presented in this thesis used pooled QC to assess intra-study variations. Indeed, pooled QCs are created from all samples analysed, and by definition contain all metabolites that could be detected in the study samples (Broadhurst et al. 2018; Dunn et al. 2011b; Dunn et al. 2011a). These pooled QC samples were used throughout the analytical experiments, but also at the very start in order to "condition" the instrumentations in order to stabilise the mass spectrometer detector response and remove retention time variability of the chromatography column (Zelena et al. 2009). In addition, an external standard was used with each injection (Leucine Enkephalin) in order to perform mass calibration of the mass spectrometer detector. Taken together, these steps enabled the acquisition of the highest quality of data.

Finally, one limitation of this thesis was that despite our rigorous selection of cases, it was highly possible that more than one sub-type of sPTB and SGA were represented in the case groups. Indeed, these two pregnancy complications still have poorly understood pathophysiology, but a consensus is that there are several pathways involved in their aetiology. That said, the samples were carefully selected to avoid presence of co-morbidities, such as gestational diabetes, and controls were matched according to age, body mass index and ethnicity. Dependencies of the altered metabolites of interest to additional confounding variables (i.e. smoking) were investigated posthoc using Wilcoxon Mann Whitney or Spearman correlation test as appropriate and showed no significant association with the altered metabolites (see **Chapter 3**, section **3.3.7**).

The findings of this thesis improved the understanding of the pathophysiology of three of the main adverse pregnancy complications: SGA, sPTB, and pre-eclampsia. In addition, this thesis further described the biological processes implicated, highlighted potential biomarkers, and demonstrated the need for more longitudinal investigations in adverse pregnancy outcome research. From these results, several experiments can be implemented in the future.

7.5. Future perspectives

Overall, clinical metabolomics offers the ability to detect subtle changes with higher sensitivity and precision than other methods, which makes it ideal for biomarker discovery and the development of precision medicine models (David S Wishart 2016). In the future, multidisciplinary research teams should aim to ensure that the clinical metabolomics field can be translated and routinely implemented in clinical settings, biomarker discovery and drug discovery.

Based on the biomarker discovery experiments for sPTB, future work could focus on biomarker validation of phosphatidylcholine (15:0/22:6) and triacylglycerol (18:3/18:2/18:3) in an independent cohort, such as the IMPROvED cohort (www.fp7-improved.eu) (Navaratnam et al. 2013). IMPROvED study aimed to recruit 5000 women in six European centres in Ireland, United Kingdom, Germany, Sweden and the Netherlands. Women recruited to IMPROvED were low-risk nulliparous women. The main outcome for this study was pre-eclampsia, however, other primary outcomes included SGA (<10^{th} customised centile), and sPTB (<37 weeks of gestation). After recruitment women were followed from the first visit (9+0 to 13+6 weeks of gestation) until birth. This study has a similar recruitment process as SCOPE study but recruited women in demographically distinct populations. This would make this cohort an ideal source of samples to undertake the validation of our potential biomarkers of SGA and sPTB, and further support our findings of altered glycerophospholipid levels associated with sPTB.

Alternatively, future work could use a multi-omics approach and analyse the samples to obtain their proteomics profiles in order to get further insight into which proteins are significantly altered between cases and controls. Multi-omics studies offer a more comprehensive and holistic approach to clinical research, which could help further the understanding of the pathophysiology and narrow down the list of potential biomarkers predictive of SGA and sPTB (Pinu et al. 2019). One such was the Maternity Log study (Sugawara et al. 2019), which aimed to follow women from initial antenatal clinic visits until the end of their pregnancy, while collecting

comprehensive multi-omics data, such as genomics, transcriptomics, proteomics, metabolomics, and including data on the microbiome. This study was only able to retain 285 participants, however, it highlighted how multi-omics clinical studies can be conducted to further understand complex pathophysiology, such as adverse pregnancy outcomes.

We reported for the first time that L-Ergothioneine may help protect mitochondrial function by reducing inflammatory response and by increasing antioxidant levels, in the RUPP rat model of pre-eclampsia. Our findings are in agreement with other work from our group performed on the same RUPP rat model, which showed that L-Ergothioneine improved hypertension, decreased antiangiogenic soluble fms-like tyrosine kinase-1 (sFlt-1), and decreased the levels of a reactive oxygen species specific to mitochondria (H_2O_2) in kidney tissue (Williamson Rachel et al. 2020). Taken together, these results showed that L-Ergothioneine has potential as a treatment for pre-eclampsia, and future work should focus on testing this in a clinical trial. As an increasing number of diseases are found to be caused by a specific metabolic process, metabolomics applications to drug discovery are not only appropriate, but offer several advantages, such as being more cost-effective than current drug development programs (David S Wishart 2016). Indeed, numerous drug targets are enzymes, for which numerous metabolites are known inhibitors. In addition, metabolomics has proven useful in reducing drug failure by performing animal toxicology screens before the start of clinical trial. The Consortium for Metabolomics Toxicology (COMET) demonstrated that metabolomics techniques are able to predict kidney and liver toxicity of lead compounds, with inexpensive, fast and non-invasive approaches (Lindon et al. 2005).

7.6. Concluding remarks

Clinical metabolomics has a lot to offer to further the understanding of disease pathophysiology, and the evidence presented here demonstrated how clinical metabolomics approaches could provide new avenues to biomarker discovery. Metabolomics approaches are especially suited to complex multi-factorial diseases,

such as the adverse pregnancy outcomes presented here. Indeed, this thesis highlighted that glycerophospholipids seem to play a key role in the pathophysiology of SGA babies and sPTB. These results showed that altered glycerophospholipids are potential biomarkers of adverse pregnancy outcomes. Although further research is needed to understand their precise associations, to determine whether they are a cause or effect of SGA and sPTB, and to validate their potential as predictive biomarker in independent cohorts, the work presented in this thesis offers further insight into the pathophysiology of SGA and sPTB.

In addition, this thesis has shown how clinical metabolomics approaches can provide invaluable information in the field of drug discovery. To date, there is no treatment for pre-eclampsia, however, this thesis has shown that the use of L-Ergothioneine treatment on the RUPP rat model of pre-eclampsia helped protect against mitochondrial oxidative stress, and demonstrated the potential of L-Ergothioneine to treat this pregnancy complication.

References

Abraham, M., Alramadhan, S., Iniguez, C., Duijts, L., Jaddoe, V. W., Den Dekker, H. T., et al. (2017). A systematic review of maternal smoking during pregnancy and fetal measurements with meta-analysis. *PLoS One, 12*(2), e0170946.

Abrahams, V. M., Bole-Aldo, P., Kim, Y. M., Straszewski-Chavez, S. L., Chaiworapongsa, T., Romero, R., et al. (2004). Divergent trophoblast responses to bacterial products mediated by TLRs. *The Journal of Immunology, 173*(7), 4286-4296.

Anandakumar, C., Chew, S., Wong, Y. C., Malarvishy, G., Po, L. U., & Ratnam, S. S. (1996). Early Asymmetric IUGR and Aneuploidy. *Journal of Obstetrics and Gynaecology Research, 22*(4), 365-370, doi:10.1111/j.1447-0756.1996.tb00990.x.

Ananth, C. V., Berkowitz, G. S., Savitz, D. A., & Lapinski, R. H. (1999). Placental abruption and adverse perinatal outcomes. [Article]. *Journal of the American Medical Association, 282*(17), 1646-1651, doi:10.1001/jama.282.17.1646.

Andrews, W. W., Hauth, J. C., Goldenberg, R. L., Gomez, R., Romero, R., & Cassell, G. H. (1995). Amniotic fluid interleukin-6: correlation with upper genital tract microbial colonization and gestational age in women delivered after spontaneous labor versus indicated delivery. *American Journal of Obstetrics and Gynecology, 173*(2), 606-612.

Antovic, A., Sennström, M., Bremme, K., & Svenungsson, E. (2018). Obstetric antiphospholipid syndrome. *Lupus science & medicine, 5*(1), e000197-e000197, doi:10.1136/lupus-2016-000197.

Arslan, M., Yazici, G., Erdem, A., Erdem, M., Ozturk Arslan, E., & Himmetoglu, O. (2004). Endothelin 1 and leptin in the pathophysiology of intrauterine growth restriction. *International Journal of Gynecology & Obstetrics, 84*(2), 120-126.

Aung, M. T., Yu, Y., Ferguson, K. K., Cantonwine, D. E., Zeng, L., McElrath, T. F., et al. (2019). Prediction and associations of preterm birth and its subtypes with eicosanoid enzymatic pathways and inflammatory markers. *Scientific reports, 9*(1), 17049, doi:10.1038/s41598-019-53448-z.

Auray-Blais, C., Raiche, E., Gagnon, R., Berthiaume, M., & Pasquier, J.-C. (2011). Metabolomics and preterm birth: What biomarkers in cervicovaginal secretions are predictive of high-risk pregnant women? *International Journal of Mass Spectrometry, 307*(1), 33-38, doi:https://doi.org/10.1016/j.ijms.2011.02.009.

Bahado-Singh, R. O., Yilmaz, A., Bisgin, H., Turkoglu, O., Kumar, P., Sherman, E., et al. (2019). Artificial intelligence and the analysis of multi-platform metabolomics data for the detection of intrauterine growth restriction. *PLoS One, 14*(4), e0214121, doi:10.1371/journal.pone.0214121.

Baig, S., Lim, J. Y., Fernandis, A. Z., Wenk, M. R., Kale, A., Su, L. L., et al. (2013). Lipidomic analysis of human placental Syncytiotrophoblast microvesicles in adverse pregnancy outcomes. *Placenta, 34*(5), 436-442, doi:https://doi.org/10.1016/j.placenta.2013.02.004.

Baker, E. J., Miles, E. A., Burdge, G. C., Yaqoob, P., & Calder, P. C. (2016). Metabolism and functional effects of plant-derived omega-3 fatty acids in humans. *Progress in Lipid Research, 64*, 30-56, doi:https://doi.org/10.1016/j.plipres.2016.07.002.

Balasuriya, C. N., Stunes, A. K., Mosti, M. P., Schei, B., Indredavik, M. S., Hals, I. K., et al. (2018). Metabolic Outcomes in Adults Born Preterm With Very Low Birthweight or Small for Gestational Age at Term: A Cohort Study. *The Journal of Clinical Endocrinology & Metabolism, 103*(12), 4437-4446.

Balayla, J., & Abenhaim, H. A. (2014). Incidence, predictors and outcomes of congenital diaphragmatic hernia: a population-based study of 32 million births in the United

States. *Journal of Maternal-Fetal & Neonatal Medicine, 27*(14), 1438-1444, doi:10.3109/14767058.2013.858691.

Barchitta, M., Maugeri, A., La Rosa, M., Magnano San Lio, R., Favara, G., Panella, M., et al. (2018). Single nucleotide polymorphisms in vitamin D receptor gene affect birth weight and the risk of preterm birth: Results from the "mamma & bambino" cohort and a meta-analysis. *Nutrients, 10*(9), 1172.

Barger, G., & Ewins, A. J. (1911). CCLVII.—The constitution of ergothioneine: a betaine related to histidine. *Journal of the Chemical Society, Transactions, 99*, 2336-2341.

Barker, D., Osmond, C., Simmonds, S., & Wield, G. (1993). The relation of small head circumference and thinness at birth to death from cardiovascular disease in adult life. *BMJ, 306*(6875), 422-426.

Bateman, R., Carruthers, R., Hoyes, J., Jones, C., Langridge, J., Millar, A., et al. (2002). A novel precursor ion discovery method on a hybrid quadrupole orthogonal acceleration time-of-flight (Q-TOF) mass spectrometer for studying protein phosphorylation. *Journal of the American Society for Mass Spectrometry, 13*(7), 792-803.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.

Berghella, V., & Saccone, G. (2016). Fetal fibronectin testing for prevention of preterm birth in singleton pregnancies with threatened preterm labor: a systematic review and metaanalysis of randomized controlled trials. *American Journal of Obstetrics and Gynecology, 215*(4), 431-438, doi:https://doi.org/10.1016/j.ajog.2016.04.038.

Berks, D., Hoedjes, M., Raat, H., Duvekot, J., Steegers, E., & Habbema, J. (2013). Risk of cardiovascular disease after pre-eclampsia and the effect of lifestyle interventions: a literature-based study. *BJOG: An International Journal of Obstetrics & Gynaecology, 120*(8), 924-931.

Beta, J., Lesmes-Heredia, C., Bedetti, C., & Akolekar, R. (2018). Risk of miscarriage following amniocentesis and chorionic villus sampling: a systematic review of the literature. *Minerva ginecologica, 70*(2), 215-219.

Birben, E., Sahiner, U. M., Sackesen, C., Erzurum, S., & Kalayci, O. (2012). Oxidative stress and antioxidant defense. *World Allergy Organization Journal, 5*(1), 9.

Bjoro, K. (1983). VASCULAR ANOMALIES OF THE UMBILICAL-CORD .1. OBSTETRIC IMPLICATIONS. *Early Human Development, 8*(2), 119-127.

Bjøro, K. (1981). Gross pathology of the placenta in intrauterine growth retardation. *Annales chirurgiae et gynaecologiae, 70*(6), 316-322.

Blencowe, H., Cousens, S., Oestergaard, M. Z., Chou, D., Moller, A.-B., Narwal, R., et al. (2012). National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *The Lancet, 379*(9832), 2162-2172, doi:https://doi.org/10.1016/S0140-6736(12)60820-4.

Bodén, R., Lundgren, M., Brandt, L., Reutfors, J., Andersen, M., & Kieler, H. (2012). Risks of adverse pregnancy and birth outcomes in women treated or not treated with mood stabilisers for bipolar disorder: population based cohort study. *BMJ : British Medical Journal, 345*, e7085, doi:10.1136/bmj.e7085.

Bodnar, L. M., Catov, J. M., Zmuda, J. M., Cooper, M. E., Parrott, M. S., Roberts, J. M., et al. (2010). Maternal serum 25-hydroxyvitamin D concentrations are associated with small-for-gestational age births in white women. *The Journal of nutrition, 140*(5), 999-1006.

Boghossian, N. S., Geraci, M., Edwards, E. M., & Horbar, J. D. (2018). Neonatal and fetal growth charts to identify preterm infants< 30 weeks gestation at risk of adverse

outcomes. *American Journal of Obstetrics and Gynecology, 219*(2), 195. e191-195. e114.

Borodina, I., Kenny, L. C., McCarthy, C. M., Paramasivan, K., Pretorius, E., Roberts, T. J., et al. (2020). The biology of ergothioneine, an antioxidant nutraceutical. *Nutrition Research Reviews, 33*(2), 190-217.

Börzsönyi, B., Demendi, C., Nagy, Z., Tóth, K., Csanád, M., Pajor, A., et al. (2011). Gene expression patterns of insulin-like growth factor 1, insulin-like growth factor 2 and insulin-like growth factor binding protein 3 in human placenta from pregnancies with intrauterine growth restriction. *Journal of perinatal medicine, 39*(6), 701-707.

Börzsönyi, B., Demendi, C., Rigó Jr, J., Szentpéteri, I., Rab, A., & Joó, J. G. (2013). The regulation of apoptosis in intrauterine growth restriction: a study of Bcl-2 and Bax gene expression in human placenta. *The Journal of Maternal-Fetal & Neonatal Medicine, 26*(4), 347-350.

Bramham, K., Parnell, B., Nelson-Piercy, C., Seed, P. T., Poston, L., & Chappell, L. C. (2014). Chronic hypertension and pregnancy outcomes: systematic review and meta-analysis. *BMJ, 348*, g2301.

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic away. *Journal of Chemometrics, 28*(4), 213-225.

Brett, K. M., Strogatz, D. S., & Savitz, D. A. (1997). Employment, job strain, and preterm delivery among women in North Carolina. *American Journal of Public Health, 87*(2), 199-204, doi:10.2105/ajph.87.2.199.

Briana, D. D., Boutsikou, M., Baka, S., Hassiakos, D., Gourgiotis, D., Marmarinos, A., et al. (2007). N-terminal parathyroid hormone-related protein levels in human intrauterine growth restricted pregnancies. *Acta obstetricia et gynecologica Scandinavica, 86*(8), 945-949.

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., et al. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics, 14*(6), 72.

Brosens, I., Pijnenborg, R., Vercruysse, L., & Romero, R. (2011). The "Great Obstetrical Syndromes" are associated with disorders of deep placentation. *American Journal of Obstetrics and Gynecology, 204*(3), 193-201.

Brown, J. M., & Hazen, S. L. (2015). The gut microbial endocrine organ: bacterially derived signals driving cardiometabolic diseases. *Annual review of medicine, 66*, 343-359.

Brown, M., Dunn, W. B., Dobson, P., Patel, Y., Winder, C., Francis-McIntyre, S., et al. (2009). Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst, 134*(7), 1322-1332.

Brown, M., Dunn, W. B., Ellis, D. I., Goodacre, R., Handl, J., Knowles, J. D., et al. (2005). A metabolome pipeline: from concept to data to knowledge. [journal article]. *Metabolomics, 1*(1), 39-51, doi:10.1007/s11306-005-1106-4.

Brown, M. A., Magee, L. A., Kenny, L. C., Karumanchi, S. A., McCarthy, F. P., Saito, S., et al. (2018). The hypertensive disorders of pregnancy: ISSHP classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertension, 13*, 291-310, doi:https://doi.org/10.1016/j.preghy.2018.05.004.

Bukowski, R., Hansen, N. I., Willinger, M., Reddy, U. M., Parker, C. B., Pinar, H., et al. (2014). Fetal Growth and Risk of Stillbirth: A Population-Based Case–Control Study. *PLOS Medicine, 11*(4), e1001633, doi:10.1371/journal.pmed.1001633.

Burton, G. J., Woods, A. W., Jauniaux, E., & Kingdom, J. C. P. (2009). Rheological and Physiological Consequences of Conversion of the Maternal Spiral Arteries for Uteroplacental Blood Flow during Human Pregnancy. *Placenta, 30*(6), 473-482, doi:https://doi.org/10.1016/j.placenta.2009.02.009.

Cai, H., Cong, W.-n., Ji, S., Rothman, S., Maudsley, S., & Martin, B. (2012). Metabolic dysfunction in Alzheimer's disease and related neurodegenerative disorders. *Current Alzheimer Research, 9*(1), 5-17.

Caliebe, J., Broekman, S., Boogaard, M., Bosch, C. A., Ruivenkamp, C. A., Oostdijk, W., et al. (2012). IGF1, IGF1R and SHOX mutation analysis in short children born small for gestational age and short children with normal birth size (idiopathic short stature). *Hormone Research in Paediatrics, 77*(4), 250-260.

Calkins, K., & Devaskar, S. U. (2011). Fetal Origins of Adult Disease. *Current Problems in Pediatric and Adolescent Health Care, 41*(6), 158-176, doi:https://doi.org/10.1016/j.cppeds.2011.01.001.

Carolan, M. (2013). Maternal age ≥45 years and maternal and perinatal outcomes: A review of the evidence. *Midwifery, 29*(5), 479-489, doi:https://doi.org/10.1016/j.midw.2012.04.001.

Carter, J. M., Demizieux, L., Campenot, R. B., Vance, D. E., & Vance, J. E. (2008). Phosphatidylcholine Biosynthesis via CTP: Phosphocholine Cytidylyltransferase β2 Facilitates Neurite Outgrowth and Branching. *Journal of Biological Chemistry, 283*(1), 202-212.

Castro-Gómez, P., Garcia-Serrano, A., Visioli, F., & Fontecha, J. (2015). Relevance of dietary glycerophospholipids and sphingolipids to human health. *Prostaglandins, leukotrienes and essential fatty acids, 101*, 41-51, doi:https://doi.org/10.1016/j.plefa.2015.07.004.

Catov, J. M., Bodnar, L. M., Kip, K. E., Hubel, C., Ness, R. B., Harger, G., et al. (2007). Early pregnancy lipid concentrations and spontaneous preterm birth. *American Journal of Obstetrics and Gynecology, 197*(6), 610.e611-610.e617, doi:https://doi.org/10.1016/j.ajog.2007.04.024.

Chalil, A., Kitson, A. P., Aristizabal Henao, J. J., Marks, K. A., Elzinga, J. L., Lamontagne-Kam, D. M. E., et al. (2018). PEMT, Δ6 desaturase, and palmitoyldocosahexaenoyl phosphatidylcholine are increased in rats during pregnancy. *Journal of lipid research, 59*(1), 123-136, doi:10.1194/jlr.M080309.

Chamorro-Garcia, R., Diaz-Castillo, C., Shoucri, B. M., Käch, H., Leavitt, R., Shioda, T., et al. (2017). Ancestral perinatal obesogen exposure results in a transgenerational thrifty phenotype in mice. *Nature communications, 8*(1), 2012.

Chaudhari, S., Otiv, M., Hoge, M., Pandit, A., & Mote, A. (2008). Growth and sexual maturation of low birth weight infants at early adolescence. *Indian pediatrics, 45*(3), 191.

Chaudhari, S., Otiv, M., Khairnar, B., Pandit, A., Hoge, M., & Sayyad, M. (2012). Pune low birth weight study—Growth from birth to adulthood. *Indian pediatrics, 49*(9), 727-732.

Cheah, I. K., & Halliwell, B. (2012). Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1822*(5), 784-793.

Cheah, I. K., Tang, R., Ye, P., Yew, T. S., Lim, K. H., & Halliwell, B. (2016). Liver ergothioneine accumulation in a guinea pig model of non-alcoholic fatty liver disease. A possible mechanism of defence? *Free radical research, 50*(1), 14-25.

Cheah, I. K., Tang, R. M., Yew, T. S., Lim, K. H., & Halliwell, B. (2017). Administration of pure ergothioneine to healthy human subjects: Uptake, metabolism, and effects on biomarkers of oxidative damage and inflammation. *Antioxidants & redox signaling, 26*(5), 193-206.

Chelbi, S. T., Wilson, M. L., Veillard, A.-C., Ingles, S. A., Zhang, J., Mondon, F., et al. (2012). Genetic and epigenetic mechanisms collaborate to control SERPINA3 expression and its association with placental diseases. *Human molecular genetics*, dds006.

Chen, T., Cao, Y., Zhang, Y., Liu, J., Bao, Y., Wang, C., et al. (2013). Random Forest in Clinical Metabolomics for Phenotypic Discrimination and Biomarker Selection. *Evidence-Based Complementary and Alternative Medicine, 2013*, 11, doi:10.1155/2013/298183.

Chen, Y.-H., Lin, H.-C., Chen, S.-F., & Lin, H.-C. (2009). Increased risk of preterm births among women with uterine leiomyoma: a nationwide population-based study. *Human reproduction, 24*(12), 3049-3056, doi:10.1093/humrep/dep320.

Cheng, Y., Leung, T., Lao, T., Chan, Y., & Sahota, D. (2016). Impact of replacing C hinese ethnicity-specific fetal biometry charts with the INTERGROWTH-21st standard. *BJOG: An International Journal of Obstetrics & Gynaecology, 123*, 48-55.

Choi, R., Choi, S., Lim, Y., Cho, Y., Kim, H., Kim, S., et al. (2016). A prospective study on serum methylmalonic acid and homocysteine in pregnant women. *Nutrients, 8*(12), 797.

Chow, L., & Lye, S. J. (1994). Expression of the gap junction protein connexin-43 is increased in the human myometrium toward term and with the onset of labor. *American Journal of Obstetrics and Gynecology, 170*(3), 788-795.

Chwalisz, K. (1994). The use of progesterone antagonists for cervical ripening and as an adjunct to labour and delivery. *Human reproduction, 9*(suppl_1), 131-161.

Cnattingius, S. (2004). The epidemiology of smoking during pregnancy: Smoking prevalence, maternal characteristics, and pregnancy outcomes. [Review]. *Nicotine and Tobacco Research, 6*(SUPPL. 2), S125-S140, doi:10.1080/14622200410001669187.

Committee on Practice Bulletins (2012). Practice Bulletin No. 130: Prediction and Prevention of Preterm Birth. *Obstetrics & Gynecology, 120*(4), 964-973, doi:10.1097/AOG.0b013e3182723b1b.

Conde-Agudelo, A., Rosas-Bermudez, A., Castaño, F., & Norton, M. H. (2012). Effects of Birth Spacing on Maternal, Perinatal, Infant, and Child Health: A Systematic Review of Causal Mechanisms. *Studies in Family Planning, 43*(2), 93-114, doi:10.1111/j.1728-4465.2012.00308.x.

Conde-Agudelo A, AT, P., SH, K., & J, V. (2013). Novel biomarkers for predicting intrauterine growth restriction: a systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics & Gynaecology, 120*(6), 681-694, doi:doi:10.1111/1471-0528.12172.

Conroy, A. L., Silver, K. L., Zhong, K., Rennie, M., Ward, P., Sarma, J. V., et al. (2013). Complement Activation and the Resulting Placental Vascular Insufficiency Drives Fetal Growth Restriction Associated with Placental Malaria. *Cell Host & Microbe, 13*(2), 215-226, doi:10.1016/j.chom.2013.01.010.

Considine, E., Thomas, G., Boulesteix, A., Khashan, A., & Kenny, L. (2018). Critical review of reporting of the data analysis step in metabolomics. *Metabolomics, 14*(1), 7.

Considine, E. C., Khashan, A. S., & Kenny, L. C. (2019). Screening for Preterm Birth: Potential for a Metabolomics Biomarker Panel. *Metabolites, 9*(5), 90.

Copper, R. L., Goldenberg, R. L., Das, A., Elder, N., Swain, M., Norman, G., et al. (1996). The preterm prediction study: Maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation. *American Journal of Obstetrics and Gynecology, 175*(5), 1286-1292.

Costet, N., Garlantézec, R., Monfort, C., Rouget, F., Gagnière, B., Chevrier, C., et al. (2011). Environmental and urinary markers of prenatal exposure to drinking water disinfection by-products, fetal growth, and duration of gestation in the PELAGIE birth cohort (Brittany, France, 2002–2006). *American Journal of Epidemiology, 175*(4), 263-275.

Cotter, D., Fahy, E., Sud, M., & Subramaniam, S. (2007). LIPID MAPS online tools for lipid research. *Nucleic acids research, 35*(suppl_2), W606-W612, doi:10.1093/nar/gkm324.

Crane, J. M. G., Van den Hof, M. C., Dodds, L., Armson, B. A., & Liston, R. (1999). Neonatal outcomes with placenta previa. *Obstetrics & Gynecology, 93*(4), 541-544, doi:https://doi.org/10.1016/S0029-7844(98)00480-3.

Cunnington, M., Kortsalioudaki, C., & Heath, P. (2013). Genitourinary pathogens and preterm birth. [Review]. *Current Opinion in Infectious Diseases, 26*(3), 219-230, doi:10.1097/QCO.0b013e328360dc31.

da Fonseca, E. B., Bittar, R. E., Carvalho, M. H., & Zugaib, M. (2003). Prophylactic administration of progesterone by vaginal suppository to reduce the incidence of spontaneous preterm birth in women at increased risk: a randomized placebo-controlled double-blind study. *American Journal of Obstetrics and Gynecology, 188*(2), 419-424.

De Onis, M., Blössner, M., & Villar, J. (1998). Levels and patterns of intrauterine growth retardation in developing countries. [Conference Paper]. *European Journal of Clinical Nutrition, 52*(SUPPL. 1), S5-S15.

de Onis, M., & Habicht, J.-P. (1996). Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee. *The American journal of clinical nutrition, 64*(4), 650-658.

Deeks, J. J., Macaskill, P., & Irwig, L. (2005). The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of clinical epidemiology, 58*(9), 882-893.

Deiana, M., Rosa, A., Casu, V., Piga, R., Assunta Dessi, M., & Aruoma, O. I. (2004). L-ergothioneine modulates oxidative damage in the kidney and liver of rats in vivo: studies upon the profile of polyunsaturated fatty acids. *Clin Nutr, 23*(2), 183-193, doi:10.1016/s0261-5614(03)00108-0.

Dekker, G. A., Lee, S. Y., North, R. A., McCowan, L. M., Simpson, N. A. B., & Roberts, C. T. (2012). Risk Factors for Preterm Birth in an International Prospective Cohort of Nulliparous Women. *PLoS One, 7*(7), e39154, doi:10.1371/journal.pone.0039154.

Delplancke, T. D., Seymour, J. V., Tong, C., Sulek, K., Xia, Y., Zhang, H., et al. (2018). Analysis of sequential hair segments reflects changes in the metabolome across the trimesters of pregnancy. *Scientific reports, 8*(1), 36.

Desai, M., ter Kuile, F. O., Nosten, F., McGready, R., Asamoa, K., Brabin, B., et al. (2007). Epidemiology and burden of malaria in pregnancy. *The Lancet Infectious Diseases, 7*(2), 93-104, doi:https://doi.org/10.1016/S1473-3099(07)70021-X.

Dessì, A., Corona, L., Pintus, R., & Fanos, V. (2018). Exposure to tobacco smoke and low birth weight: from epidemiology to metabolomics. *Expert review of proteomics, 15*(8), 647-656.

Dessì, 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, doi:https://doi.org/10.1016/j.bpobgyn.2014.04.023.

Dessì, A., Ottonello, G., & Fanos, V. (2012). Physiopathology of intrauterine growth retardation: from classic data to metabolomics. *The Journal of Maternal-Fetal & Neonatal Medicine, 25*(sup5), 13-18, doi:10.3109/14767058.2012.714639.

Dhobale, M. V., Wadhwani, N., Mehendale, S. S., Pisal, H. R., & Joshi, S. R. (2011). Reduced levels of placental long chain polyunsaturated fatty acids in preterm deliveries. *Prostaglandins, leukotrienes and essential fatty acids, 85*(3-4), 149-153.

Di Giulio, A., Carelli, S., Castoldi, R., Gorio, A., Taricco, E., & Cetin, I. (2004). Plasma amino acid concentrations throughout normal pregnancy and early stages of intrauterine growth restricted pregnancy. *The Journal of Maternal-Fetal & Neonatal Medicine, 15*(6), 356-362.

Diderholm, B. (2009). Perinatal energy metabolism with reference to IUGR & SGA: studies in pregnant women & newborn infants. *Indian J Med Res, 130*(5), 612-617.

Djakovic, A., Rieger, L., Wirbelauer, J., Kalla, J., & Dietl, J. (2007). [Severe foetal growth retardation in a patient with uterus bicornis, velamentous insertion and partial placental abruption in the 26th week of gestation--a case report]. *Zeitschrift fur Geburtshilfe und Neonatologie, 211*(4), 169-173, doi:10.1055/s-2007-981219.

dos Santos-Weiss, I. C., Réa, R. R., Fadel-Picheth, C. M., Rego, F. G., Pedrosa, F. d. O., Gillery, P., et al. (2013). The plasma logarithm of the triglyceride/HDL-cholesterol ratio is a predictor of low risk gestational diabetes in early pregnancy. *Clinica Chimica Acta, 418*, 1-4.

Draper, J., Enot, D. P., Parker, D., Beckmann, M., Snowdon, S., Lin, W., et al. (2009). Metabolite signal identification in accurate mass metabolomics data with MZedDB, an interactive m/z annotation tool utilising predicted ionisation behaviour'rules'. *BMC Bioinformatics, 10*(1), 227.

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., et al. (2011a). 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., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011b). 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., Erban, A., Weber, R. J., Creek, D. J., Brown, M., Breitling, R., et al. (2013). Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics, 9*(1), 44-66.

Duttaroy, A. K. (2009). Transport of fatty acids across the human placenta: A review. *Progress in Lipid Research, 48*(1), 52-61, doi:https://doi.org/10.1016/j.plipres.2008.11.001.

Ertl, R., Christina, K., Samaha, R., Akolekar, R., & Nicolaides, K. H. (2012). Maternal serum vitamin D at 11–13 weeks in pregnancies delivering small for gestational age neonates. *Fetal Diagnosis and Therapy, 31*(2), 103-108.

Everett, T. R., Wilkinson, I. B., & Lees, C. C. (2012). Drug development in preeclampsia: a 'no go'area? *The Journal of Maternal-Fetal & Neonatal Medicine, 25*(1), 50-52.

Fang, J., Madhavan, S., & Alderman, M. H. (1999). Low birth weight: Race and maternal nativity - Impact of community income. *Pediatrics, 103*(1), doi:10.1542/peds.103.1.e5.

Favretto, D., Cosmi, E., Ragazzi, E., Visentin, S., Tucci, M., Fais, P., et al. (2012). Cord blood metabolomic profiling in intrauterine growth restriction. *Analytical and bioanalytical chemistry, 402*(3), 1109-1121.

Fazeli, A., Bruce, C., & Anumba, D. (2005). Characterization of Toll-like receptors in the female reproductive tract in humans. *Human reproduction, 20*(5), 1372-1378.

Ferrero, D. M., Larson, J., Jacobsson, B., Di Renzo, G. C., Norman, J. E., Martin, J. N., Jr., et al. (2016). Cross-Country Individual Participant Analysis of 4.1 Million Singleton Births in 5 Countries with Very High Human Development Index Confirms Known Associations but Provides No Biologic Explanation for 2/3 of All Preterm Births. *PLoS One, 11*(9), e0162506, doi:10.1371/journal.pone.0162506.

Fidel Jr, P. L., Romero, R., Maymon, E., & Hertelendy, F. (1998). Bacteria-induced or bacterial product-induced preterm parturition in mice and rabbits is preceded by a significant fall in serum progesterone concentrations. *The Journal of Maternal-Fetal Medicine, 7*(5), 222-226.

Fiehn, O. (2002). Metabolomics — the link between genotypes and phenotypes. In C. Town (Ed.), *Functional Genomics* (pp. 155-171). Dordrecht: Springer Netherlands.

Fiehn, O., Robertson, D., Griffin, J., van der Werf, M., Nikolau, B., Morrison, N., et al. (2007). The metabolomics standards initiative (MSI). *Metabolomics, 3*(3), 175-178.

Fieni, S., & Gramellini, D. (2004). Very-early-onset discordant growth in monochorionic twin pregnancy. *Obstetrics and gynecology, 103*(5), 1115-1117, doi:10.1097/01.AOG.0000125872.96052.9f.

Figueras, F., & Gratacos, E. (2014). Update on the diagnosis and classification of fetal growth restriction and proposal of a stage-based management protocol. *Fetal Diagnosis and Therapy, 36*(2), 86-98.

Fiscella, K. (1996). Race, Perinatal Outcome, and Amniotic Infection. *Obstetrical & Gynecological Survey, 51*(1), 60-66.

Fischer, S. T., Lili, L. N., Li, S., Tran, V. T., Stewart, K. B., Schwartz, C. E., et al. (2017). Low-level maternal exposure to nicotine associates with significant metabolic perturbations in second-trimester amniotic fluid. *Environment International, 107*, 227-234.

Fisk, N., Ronderos-Dumit, D., Tannirandorn, Y., Nicolini, U., Talbert, D., & Rodeck, C. (1992). Normal amniotic pressure throughout gestation. *BJOG: An International Journal of Obstetrics & Gynaecology, 99*(1), 18-22.

Florio, P., Marinoni, E., Di Iorio, R., Bashir, M., Ciotti, S., Sacchi, R., et al. (2006). Urinary S100B protein concentrations are increased in intrauterine growth-retarded newborns. *Pediatrics, 118*(3), e747-e754.

Forray, A. (2016). Substance use during pregnancy. *F1000Research, 5*, F1000 Faculty Rev-1887, doi:10.12688/f1000research.7645.1.

Forster, R., Spézia, F., Papineau, D., Sabadie, C., Erdelmeier, I., Moutet, M., et al. (2015). Reproductive safety evaluation of L-Ergothioneine. *Food and Chemical Toxicology, 80*, 85-91, doi:https://doi.org/10.1016/j.fct.2015.02.019.

Fox, N. S., Roman, A. S., Stern, E. M., Gerber, R. S., Saltzman, D. H., & Rebarber, A. (2014). Type of congenital uterine anomaly and adverse pregnancy outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine, 27*(9), 949-953, doi:10.3109/14767058.2013.847082.

Franzoni, F., Colognato, R., Galetta, F., Laurenza, I., Barsotti, M., Di Stefano, R., et al. (2006). An in vitro study on the free radical scavenging capacity of ergothioneine: comparison with reduced glutathione, uric acid and trolox. *Biomedicine & pharmacotherapy, 60*(8), 453-457.

Frolkis, A., Knox, C., Lim, E., Jewison, T., Law, V., Hau, D. D., et al. (2010). SMPDB: The Small Molecule Pathway Database. *Nucleic acids research, 38*(Database issue), D480-D487, doi:10.1093/nar/gkp1002.

Funke, C., Schneider, S. A., Berg, D., & Kell, D. B. (2013). Genetics and iron in the systems biology of Parkinson's disease and some related disorders. *Neurochemistry international, 62*(5), 637-652.

Galan, H. L., & Grobman, W. (2015). Fetal growth restriction. *Protocols for High-Risk Pregnancies*, 340.

Galinsky, R., Polglase, G. R., Hooper, S. B., Black, M. J., & Moss, T. J. M. (2013). The Consequences of Chorioamnionitis: Preterm Birth and Effects on Development. *Journal of pregnancy, 2013*, 11, doi:10.1155/2013/412831.

Garcia, D., & Erkan, D. (2018). Diagnosis and management of the antiphospholipid syndrome. *New England Journal of Medicine, 378*(21), 2010-2021.

Gascoin-Lachambre, G., Buffat, C., Rebourcet, R., Chelbi, S., Rigourd, V., Mondon, F., et al. (2010). Cullins in human intra-uterine growth restriction: expressional and epigenetic alterations. *Placenta, 31*(2), 151-157.

Germain, S. J., Sacks, G. P., Soorana, S. R., Sargent, I. L., & Redman, C. W. (2007). Systemic inflammatory priming in normal pregnancy and preeclampsia: the role of circulating syncytiotrophoblast microparticles. *The Journal of Immunology, 178*(9), 5949-5956.

Gernand, A. D., Simhan, H. N., Klebanoff, M. A., & Bodnar, L. M. (2013). Maternal serum 25-hydroxyvitamin D and measures of newborn and placental weight in a US

multicenter cohort study. *The Journal of Clinical Endocrinology & Metabolism, 98*(1), 398-404.

Gika, H. G., Theodoridis, G. A., & Wilson, I. D. (2008). Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabonomic analysis of Zucker rat urine. *Journal of separation science, 31*(9), 1598-1608.

Goldenberg, R. L., Cliver, S. P., Mulvihill, F. X., Hickey, C. A., Hoffman, H. J., Klerman, L. V., et al. (1996). Medical, psychosocial, and behavioral risk factors do not explain the increased risk for low birth weight among black women. *American Journal of Obstetrics and Gynecology, 175*(5), 1317-1324, doi:https://doi.org/10.1016/S0002-9378(96)70048-0.

Goldenberg, R. L., Culhane, J. F., Iams, J. D., & Romero, R. (2008). Epidemiology and causes of preterm birth. *The Lancet, 371*(9606), 75-84, doi:https://doi.org/10.1016/S0140-6736(08)60074-4.

Goldenberg, R. L., Goepfert, A. R., & Ramsey, P. S. (2005). Biochemical markers for the prediction of preterm birth. *American Journal of Obstetrics and Gynecology, 192*(5), S36-S46.

Goldenberg, R. L., Hauth, J. C., & Andrews, W. W. (2000). Intrauterine infection and preterm delivery. *New England Journal of Medicine, 342*(20), 1500-1507.

Goldenberg, R. L., Mercer, B. M., Meis, P. J., Copper, R. L., Das, A., McNellis, D., et al. (1996). The preterm prediction study: fetal fibronectin testing and spontaneous preterm birth. *Obstetrics & Gynecology, 87*(5), 643-648.

Gonçalves, L. F., Chaiworapongsa, T., & Romero, R. (2002). Intrauterine infection and prematurity. *Mental retardation and developmental disabilities research reviews, 8*(1), 3-13.

Gong, S., Sovio, U., Aye, I. L., Gaccioli, F., Dopierala, J., Johnson, M. D., et al. (2018). Placental polyamine metabolism differs by fetal sex, fetal growth restriction, and preeclampsia. *JCI insight, 3*(13).

Goodacre, R., Broadhurst, D., Smilde, A. K., Kristal, B. S., Baker, J. D., Beger, R., et al. (2007). Proposed minimum reporting standards for data analysis in metabolomics. [journal article]. *Metabolomics, 3*(3), 231-241, doi:10.1007/s11306-007-0081-3.

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.

Gordijn, S., Beune, I., Thilaganathan, B., Papageorghiou, A., Baschat, A., Baker, P., et al. (2016). Consensus definition of fetal growth restriction: a Delphi procedure. *Ultrasound in Obstetrics & Gynecology, 48*(3), 333-339.

Gorodeski, I. G., Geier, A., Lunenfeld, B., Beery, R., & Bahary, C. M. (1987). Progesterone (P) receptor dynamics in estrogen primed normal human cervix following P injection. *Fertility and sterility, 47*(1), 108-113.

Goss, K. C. W., Goss, V. M., Townsend, J. P., Koster, G., Clark, H. W., & Postle, A. D. (2020). Postnatal adaptations of phosphatidylcholine metabolism in extremely preterm infants: implications for choline and PUFA metabolism. *The American journal of clinical nutrition, 112*(6), 1438-1447, doi:10.1093/ajcn/nqaa207.

Goto, E. (2016). Maternal anthropometry to predict small for gestational age: a meta-analysis. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 203*, 193-198.

Graham, S., Chevallier, O., Kumar, P., Türkoğlu, O., & Bahado-Singh, R. (2017). Metabolomic profiling of brain from infants who died from Sudden Infant Death Syndrome reveals novel predictive biomarkers. *Journal of Perinatology, 37*(1), 91-97.

Grandone, E., Colaizzo, D., Vecchione, G., Sciannamè, N., Notarangelo, A., Croce, A. I., et al. (2006). Homocysteine levels in amniotic fluid. *Thrombosis and haemostasis, 95*(04), 625-628.

Grandone, E., Margaglione, M., Colaizzo, D., Pavone, G., Paladini, D., Martinelli, P., et al. (2002). Lower birth-weight in neonates of mothers carrying factor V G1691A and factor II A (20210) mutations. *Haematologica, 87*(2), 177-181.

Granger, J. P., LaMarca, B. B., Cockrell, K., Sedeek, M., Balzi, C., Chandler, D., et al. (2006). Reduced uterine perfusion pressure (RUPP) model for studying cardiovascular-renal dysfunction in response to placental ischemia. *Methods Mol Med, 122*, 383-392.

Grantham-McGregor, S. (1998). Small for gestational age, term babies, in the first six years of life. *European Journal of Clinical Nutrition, 52*, S59-64.

Gray, N., Zia, R., King, A., Patel, V. C., Wendon, J., McPhail, M. J., et al. (2017). High-speed quantitative UPLC-MS analysis of multiple amines in human plasma and serum via precolumn derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate: application to acetaminophen-induced liver failure. *Analytical Chemistry, 89*(4), 2478-2487.

Gremlich, S., Damnon, F., Reymondin, D., Braissant, O., Schittny, J., Baud, D., et al. (2014). The long non-coding RNA NEAT1 is increased in IUGR placentas, leading to potential new hypotheses of IUGR origin/development. *Placenta, 35*(1), 44-49.

Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., & Deigner, H.-P. (2010). Targeted Metabolomics for Biomarker Discovery. *Angewandte Chemie International Edition, 49*(32), 5426-5445, doi:10.1002/anie.200905579.

Guang, B. W., Monique, N. A., Hope, W., Line, L., Christina, S., & Qin, W. S. (2018). Association between vitamin D supplementation during pregnancy and offspring growth, morbidity, and mortality: A systematic review and meta-analysis. *JAMA pediatrics, 172*(7), 635-645.

Haahr, T., Ersbøll, A. S., Karlsen, M. A., Svare, J., Sneider, K., Hee, L., et al. (2016). Treatment of bacterial vaginosis in pregnancy in order to reduce the risk of spontaneous preterm delivery – a clinical recommendation. *Acta obstetricia et gynecologica Scandinavica, 95*(8), 850-860, doi:10.1111/aogs.12933.

Haggarty, P., Ashton, J., Joynson, M., Abramovich, D., & Page, K. (1999). Effect of maternal polyunsaturated fatty acid concentration on transport by the human placenta. *Neonatology, 75*(6), 350-359.

Hales, C. N., & Barker, D. J. (2001). The thrifty phenotype hypothesis: Type 2 diabetes. *British Medical Bulletin, 60*(1), 5-20.

Halliwell, B., Cheah, I. K., & Tang, R. M. Y. (2018). Ergothioneine – a diet-derived antioxidant with therapeutic potential. *FEBS Letters, 592*(20), 3357-3366, doi:10.1002/1873-3468.13123.

Heazell, A. E. P., Newman, L., Lean, S. C., & Jones, R. L. (2018). Pregnancy outcome in mothers over the age of 35. *Current Opinion in Obstetrics and Gynecology, 30*(6), 337-343, doi:10.1097/gco.0000000000000494.

Heinonen, S., Taipale, P., & Saarikoski, S. (2001). Weights of placentae from small-for-gestational age infants revisited. *Placenta, 22*(5), 399-404, doi:10.1053/plac.2001.0630.

Hendler, I., Goldenberg, R. L., Mercer, B. M., Iams, J. D., Meis, P. J., Moawad, A. H., et al. (2005). The Preterm Prediction study: Association between maternal body mass index and spontaneous and indicated preterm birth. *American Journal of Obstetrics and Gynecology, 192*(3), 882-886, doi:https://doi.org/10.1016/j.ajog.2004.09.021.

Hendrix, N., & Berghella, V. (2008). Non-placental causes of intrauterine growth restriction. *Seminars in Perinatology, 32*(3), 161-165, doi:10.1053/j.semperi.2008.02.004.

Hitti, J., Riley, D. E., Krohn, M. A., Hillier, S. L., Agnew, K. J., Krieger, J. N., et al. (1997). Broad-spectrum bacterial rDNA polymerase chain reaction assay for detecting amniotic fluid infection among women in premature labor. *Clinical infectious diseases, 24*(6), 1228-1232.

Hobel, C. J. (2004). Stress and preterm birth. *Clinical obstetrics and gynecology, 47*(4), 856-880.

Hobel, C. J., Dunkel-Schetter, C., Roesch, S. C., Castro, L. C., & Arora, C. P. (1999). Maternal plasma corticotropin-releasing hormone associated with stress at 20 weeks' gestation in pregnancies ending in preterm delivery. *American Journal of Obstetrics and Gynecology, 180*(1), S257-S263.

Hogeveen, M., Blom, H. J., & den Heijer, M. (2011). Maternal homocysteine and small-for-gestational-age offspring: systematic review and meta-analysis. *The American journal of clinical nutrition, 95*(1), 130-136, doi:10.3945/ajcn.111.016212.

Holick, M. F., Binkley, N. C., Bischoff-Ferrari, H. A., Gordon, C. M., Hanley, D. A., Heaney, R. P., et al. (2011). Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. *The Journal of Clinical Endocrinology & Metabolism, 96*(7), 1911-1930.

Honest, H., Bachmann, L. M., Gupta, J. K., Kleijnen, J., & Khan, K. S. (2002). Accuracy of cervicovaginal fetal fibronectin test in predicting risk of spontaneous preterm birth: systematic review. *BMJ, 325*(7359), 301, doi:10.1136/bmj.325.7359.301.

Horgan, R. P., Broadhurst, D. I., Walsh, S. K., Dunn, W. B., Brown, M., Roberts, C. T., et al. (2011). Metabolic profiling uncovers a phenotypic signature of small for gestational age in early pregnancy. *Journal of proteome research, 10*(8), 3660-3673.

Huang, L., Shen, Z., Xu, Q., Huang, X., Chen, Q., & Li, D. (2013). Increased levels of microRNA-424 are associated with the pathogenesis of fetal growth restriction. *Placenta, 34*(7), 624-627, doi:http://dx.doi.org/10.1016/j.placenta.2013.04.009.

Huang, Q.-t., Huang, Q., Zhong, M., Wei, S.-s., Luo, W., Li, F., et al. (2015). Chronic hepatitis C virus infection is associated with increased risk of preterm birth: a meta-analysis of observational studies. *Journal of Viral Hepatitis, 22*(12), 1033-1042, doi:10.1111/jvh.12430.

Hubel, C. A., Shakir, Y., Gallaher, M. J., McLaughlin, M. K., & Roberts, J. M. (1998). Low-density lipoprotein particle size decreases during normal pregnancy in association with triglyceride increases. *Journal of the Society for Gynecologic Investigation, 5*(5), 244-250.

Hug L., Sharrow D., & You D. (2017). Levels & trends in child mortality: report 2017. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation. *New York: United Nations Children's Fund.*

Huppertz, B., Hemmings, D., Renaud, S., Bulmer, J., Dash, P., & Chamley, L. (2005). Extravillous trophoblast apoptosis–a workshop report. *Placenta, 26*, S46-S48.

Iams, J., Newman, R., Thom, E., Goldenberg, R., Mueller-Heubach, E., Moawad, A., et al. (2002). National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units. Frequency of uterine contractions and the risk of spontaneous preterm delivery. *N Engl J Med, 346*(4), 250-255.

Iams, J. D., Johnson, F. F., Sonck, J., Sachs, L., Gebauer, C., & Samuels, P. (1995). Cervical competence as a continuum: a study of ultrasonographic cervical length and obstetric performance. *American Journal of Obstetrics and Gynecology, 172*(4), 1097-1106.

Iannucci, T. A., Tomich, P. G., & Gianopoulos, J. G. (1996). Etiology and outcome of extremely low-birth-weight infants. *American Journal of Obstetrics and Gynecology, 174*(6), 1896-1902, doi:https://doi.org/10.1016/S0002-9378(96)70226-0.

Imbard, A., Blom, H. J., Schlemmer, D., Barto, R., Czerkiewicz, I., Rigal, O., et al. (2013). Methylation metabolites in amniotic fluid depend on gestational age. *Prenatal Diagnosis, 33*(9), 848-855, doi:10.1002/pd.4142.

J., G., M., M., M., W., & A., C. (1995). An adjustable fetal weight standard. *Ultrasound in Obstetrics & Gynecology, 6*(3), 168-174, doi:doi:10.1046/j.1469-0705.1995.06030168.x.

Jansson, T., & Powell, T. L. (2007). Role of the placenta in fetal programming: underlying mechanisms and potential interventional approaches. *Clinical science, 113*(1), 1-13.

Jebbink, J., Wolters, A., Fernando, F., Afink, G., van der Post, J., & Ris-Stalpers, C. (2012). Molecular genetics of preeclampsia and HELLP syndrome—a review. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1822*(12), 1960-1969.

Jewison, T., Su, Y., Disfany, F. M., Liang, Y., Knox, C., Maciejewski, A., et al. (2014). SMPDB 2.0: big improvements to the Small Molecule Pathway Database. *Nucleic acids research, 42*(Database issue), D478-D484, doi:10.1093/nar/gkt1067.

Jiang, S., Jiang, J., Xu, H., Wang, S., Liu, Z., Li, M., et al. (2017). Maternal dyslipidemia during pregnancy may increase the risk of preterm birth: A meta-analysis. *Taiwanese Journal of Obstetrics and Gynecology, 56*(1), 9-15, doi:https://doi.org/10.1016/j.tjog.2016.07.012.

Jolliffe, I. T. (1990). Principal component analysis: a beginner's guide—I. Introduction and application. *Weather, 45*(10), 375-382.

Jones, H. N., Powell, T. L., & Jansson, T. (2007). Regulation of Placental Nutrient Transport – A Review. *Placenta, 28*(8), 763-774, doi:https://doi.org/10.1016/j.placenta.2007.05.002.

Jones, K. L., Mansell, A., Patella, S., Scott, B. J., Hedger, M. P., de Kretser, D. M., et al. (2007). Activin A is a critical component of the inflammatory response, and its binding protein, follistatin, reduces mortality in endotoxemia. *Proceedings of the National Academy of Sciences, 104*(41), 16239-16244.

Jones, S., & Challis, J. (1989). Local stimulation of prostagiandin production by corttcotropin-releasing hormone in human fetal membranes and placenta. *Biochemical and biophysical research communications, 159*(1), 192-199.

Jung, E. Y., Park, J. W., Ryu, A., Lee, S. Y., Cho, S.-h., & Park, K. H. (2016). Prediction of impending preterm delivery based on sonographic cervical length and different cytokine levels in cervicovaginal fluid in preterm labor. *Journal of Obstetrics and Gynaecology Research, 42*(2), 158-165, doi:10.1111/jog.12882.

Kammers, K., Cole, R. N., Tiengwe, C., & Ruczinski, I. (2015). Detecting significant changes in protein abundance. *EuPA open proteomics, 7*, 11-19.

Kanayama, N., & Fukamizu, H. (1989). Mechanical stretching increases prostaglandin E2 in cultured human amnion cells. *Gynecologic and obstetric investigation, 28*(3), 123-126.

Kar, S., Wong, M., Rogozinska, E., & Thangaratinam, S. (2016). Effects of omega-3 fatty acids in prevention of early preterm delivery: a systematic review and meta-analysis of randomized studies. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 198*, 40-46, doi:https://doi.org/10.1016/j.ejogrb.2015.11.033.

Karalis, K., Sano, H., Redwine, J., Listwak, S., Wilder, R. L., & Chrousos, G. P. (1991). Autocrine or paracrine inflammatory actions of corticotropin-releasing hormone in vivo. *Science, 254*(5030), 421-423.

Katz, J., Lee, A. C., Kozuki, N., Lawn, J. E., Cousens, S., Blencowe, H., et al. (2013). Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *The Lancet, 382*(9890), 417-425.

Katz, J., Wu, L. A., Mullany, L. C., Coles, C. L., Lee, A. C., Kozuki, N., et al. (2014). Prevalence of small-for-gestational-age and its mortality risk varies by choice of birth-weight-for-gestation reference population. *PLoS One, 9*(3), e92074.

Kawano, H., Otani, M., Takeyama, K., Kawai, Y., Mayumi, T., & Hama, T. (1982). Studies on ergothioneine. VI. Distribution and fluctuations of ergothioneine in rats. *Chemical and pharmaceutical bulletin, 30*(5), 1760-1765.

Kawashima, Y., Higaki, K., Fukushima, T., Hakuno, F., Nagaishi, J. i., Hanaki, K., et al. (2012). Novel missense mutation in the IGF-I receptor L2 domain results in intrauterine and postnatal growth retardation. *Clinical endocrinology, 77*(2), 246-254.

KEGG. https://www.genome.jp/kegg/.

Kyoto encyclopedia of genes and genomes, . https://www.genome.jp/kegg/ Accessed Acessed 20 Dec 2018.

Kell, D. B., & Kenny, L. C. (2016). A Dormant Microbial Component in the Development of Preeclampsia. [Review]. *Frontiers in medicine, 3*(60), doi:10.3389/fmed.2016.00060.

Kelly, R., Leask, R., & Calder, A. (1992). Choriodecidual production of interleukin-8 and mechanism of parturition. *The Lancet, 339*(8796), 776-777.

Kenny, L. C., Black, M. A., Poston, L., Taylor, R., Myers, J. E., Baker, P. N., et al. (2014). Early Pregnancy Prediction of Preeclampsia in Nulliparous Women, Combining Clinical Risk and Biomarkers The Screening for Pregnancy Endpoints (SCOPE) International Cohort Study. *Hypertension, 64*(3), 644-652.

Kenny, L. C., Black, M. A., Poston, L., Taylor, R., Myers, J. E., Baker, P. N., et al. (2014). Early pregnancy prediction of preeclampsia in nulliparous women, combining clinical risk and biomarkers: The Screening for Pregnancy Endpoints (SCOPE) international cohort study. [Article]. *Hypertension, 64*(3), 644-652, doi:10.1161/HYPERTENSIONAHA.114.03578.

Kenny, L. C., Broadhurst, D. I., Dunn, W., Brown, M., North, R. A., McCowan, L., et al. (2010). Robust Early Pregnancy Prediction of Later Preeclampsia Using Metabolomic Biomarkers. *Hypertension, 56*(4), 741-749, doi:10.1161/hypertensionaha.110.157297.

Kenny, L. C., & Kell, D. B. (2018). Immunological tolerance, pregnancy, and preeclampsia: the roles of semen microbes and the father. *Frontiers in medicine, 4*, 239.

Kesmodel, U., Olsen, S. F., & Secher, N. J. (2000). Does alcohol increase the risk of preterm delivery? [Article]. *Epidemiology, 11*(5), 512-518, doi:10.1097/00001648-200009000-00005.

Khamis, M. M., Adamko, D. J., & El-Aneed, A. (2020). STRATEGIES AND CHALLENGES IN METHOD DEVELOPMENT AND VALIDATION FOR THE ABSOLUTE QUANTIFICATION OF ENDOGENOUS BIOMARKER METABOLITES USING LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY. *Mass spectrometry reviews, n/a*(n/a), doi:10.1002/mas.21607.

Khashan, A. S., Evans, M., Kublickas, M., McCarthy, F. P., Kenny, L. C., Stenvinkel, P., et al. (2019). Preeclampsia and risk of end stage kidney disease: A Swedish nationwide cohort study. *PLOS Medicine, 16*(7), e1002875, doi:10.1371/journal.pmed.1002875.

Khoury, M. J., Erickson, J. D., Cordero, J. F., & McCarthy, B. J. (1988). CONGENITAL-MALFORMATIONS AND INTRAUTERINE GROWTH-RETARDATION - A POPULATION STUDY. *Pediatrics, 82*(1), 83-90.

Kiely, M. E., Zhang, J. Y., Kinsella, M., Khashan, A. S., & Kenny, L. C. (2016). Vitamin D status is associated with uteroplacental dysfunction indicated by pre-eclampsia and small-for-gestational-age birth in a large prospective pregnancy cohort in Ireland with low vitamin D status. *The American journal of clinical nutrition, 104*(2), 354-361.

Kim, J., Kim, H., Roh, H., & Kwon, Y. (2018). Causes of hyperhomocysteinemia and its pathological significance. *Archives of pharmacal research, 41*(4), 372-383.

Kim, Y. M., Romero, R., Chaiworapongsa, T., Kim, G. J., Kim, M. R., Kuivaniemi, H., et al. (2004). Toll-like receptor-2 and-4 in the chorioamniotic membranes in spontaneous labor at term and in preterm parturition that are associated with chorioamnionitis. *American Journal of Obstetrics and Gynecology, 191*(4), 1346-1355.

Kiserud, T., Piaggio, G., Carroli, G., Widmer, M., Carvalho, J., Neerup Jensen, L., et al. (2017). The World Health Organization Fetal Growth Charts: A Multinational Longitudinal Study of Ultrasound Biometric Measurements and Estimated Fetal Weight. *PLOS Medicine, 14*(1), e1002220, doi:10.1371/journal.pmed.1002220.

Kleinrouweler, C. E., Wiegerinck, M. M., Ris-Stalpers, C., Bossuyt, P. M., van der Post, J. A., von Dadelszen, P., et al. (2012). Accuracy of circulating placental growth factor, vascular endothelial growth factor, soluble fms-like tyrosine kinase 1 and soluble endoglin in the prediction of pre-eclampsia: a systematic review and meta-analysis. *BJOG: An International Journal of Obstetrics & Gynaecology, 119*(7), 778-787.

Kloeck, F., & Jung, H. (1973). In vitro release of prostaglandins from the human myometrium under the influence of stretching. *American Journal of Obstetrics and Gynecology, 115*(8), 1066-1069.

Korzeniewski, S. J., Allred, E. N., Joseph, R. M., Heeren, T., Kuban, K. C., O'Shea, T. M., et al. (2017). Neurodevelopment at age 10 years of children born< 28 weeks with fetal growth restriction. *Pediatrics, 140*(5), e20170697.

Kozuki, N., Lee, A. C. C., Silveira, M. F., Sania, A., Vogel, J. P., Adair, L., et al. (2013). The associations of parity and maternal age with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis. *BMC Public Health, 13*(3), S2, doi:10.1186/1471-2458-13-S3-S2.

Kramer, M. S., Papageorghiou, A., Culhane, J., Bhutta, Z., Goldenberg, R. L., Gravett, M., et al. (2012). Challenges in defining and classifying the preterm birth syndrome. *American Journal of Obstetrics and Gynecology, 206*(2), 108-112, doi:https://doi.org/10.1016/j.ajog.2011.10.864.

Krampl, E. (2002). Pregnancy at high altitude. *Ultrasound in Obstetrics & Gynecology, 19*(6), 535-539, doi:10.1046/j.1469-0705.2002.00738.x.

Lager, S., & Powell, T. L. (2012). Regulation of nutrient transport across the placenta. *Journal of pregnancy, 2012*.

Lappen, J. R., & Myers, S. A. (2017). The systematic error in the estimation of fetal weight and the underestimation of fetal growth restriction. *American Journal of Obstetrics and Gynecology, 216*(5), 477-483, doi:https://doi.org/10.1016/j.ajog.2017.02.013.

Laudański, T., & Rocki, W. (1975). The effects on stretching and prostaglandin F2alpha on the contractile and bioelectric activity of the uterus in rat. *Acta Physiologica Polonica, 26*(4), 385-393.

Leal, A. C., Montenegro, L. R., Saito, R. F., Ribeiro, T. C., Coutinho, D. C., Mendonca, B. B., et al. (2013). Analysis of the insulin-like growth factor 1 receptor gene in children born small for gestational age: in vitro characterization of a novel mutation (p. Arg511Trp). *Clinical endocrinology, 78*(4), 558-563.

Leal, A. d. C., Canton, A. P. M., Montenegro, L. R., Coutinho, D. C., Arnhold, I. J. P., & Jorge, A. A. d. L. (2011). Mutations in insulin-like growth factor receptor 1 gene (IGF1R) resulting in intrauterine and postnatal growth retardation. *Arquivos Brasileiros de Endocrinologia & Metabologia, 55*(8), 541-549.

Lee, A. C., Katz, J., Blencowe, H., Cousens, S., Kozuki, N., Vogel, J. P., et al. (2013). National and regional estimates of term and preterm babies born small for gestational age in 138 low-income and middle-income countries in 2010. *The Lancet Global Health, 1*(1), e26-e36.

Lee, H.-C., & Yokomizo, T. (2018). Applications of mass spectrometry-based targeted and non-targeted lipidomics. *Biochemical and biophysical research communications, 504*(3), 576-581.

Lee, M. H., Jeon, Y. J., Lee, S. M., Park, M. H., Jung, S.-C., & Kim, Y. J. (2010). Placental gene expression is related to glucose metabolism and fetal cord blood levels of insulin and insulin-like growth factors in intrauterine growth restriction. *Early Human Development, 86*(1), 45-50.

Lee, P. A., Chernausek, S. D., Hokken-Koelega, A. C. S., & Czernichow, P. (2003). International Small for Gestational Age Advisory Board Consensus Development Conference Statement: Management of Short Children Born Small for Gestational Age, April 24–October 1, 2001. *Pediatrics, 111*(6), 1253-1261.

Lees, C., Marlow, N., Arabin, B., Bilardo, C. M., Brezinka, C., Derks, J., et al. (2013). Perinatal morbidity and mortality in early-onset fetal growth restriction: cohort outcomes of the trial of randomized umbilical and fetal flow in Europe (TRUFFLE). *Ultrasound in Obstetrics & Gynecology, 42*(4), 400-408.

Leffelaar, E. R., Vrijkotte, T. G., & van Eijsden, M. (2010). Maternal early pregnancy vitamin D status in relation to fetal and neonatal growth: results of the multi-ethnic Amsterdam Born Children and their Development cohort. *British Journal of Nutrition, 104*(1), 108-117.

Leite, D. F. B., Morillon, A.-C., Melo Júnior, E. F., Souza, R. T., Khashan, A. S., Baker, P. N., et al. (2018). Metabolomics for predicting fetal growth restriction: protocol for a systematic review and meta-analysis. *BMJ Open, 8*(12), e022743, doi:10.1136/bmjopen-2018-022743.

Leite, D. F. B., Morillon, A.-C., Melo Júnior, E. F., Souza, R. T., McCarthy, F. P., Khashan, A., et al. (2019). Examining the predictive accuracy of metabolomics for small-for-gestational-age babies: a systematic review. *BMJ Open, 9*(8), e031238, doi:10.1136/bmjopen-2019-031238.

Lenaerts, C., Bond, L., Tuytten, R., & Blankert, B. (2018). Revealing of endogenous Marinobufagin by an ultra-specific and sensitive UHPLC-MS/MS assay in pregnant women. *Talanta, 187*, 193-199, doi:https://doi.org/10.1016/j.talanta.2018.05.020.

Lewi, L., Gucciardo, L., Huber, A., Jani, J., Van Mieghem, T., Done, E., et al. (2008). Clinical outcome and placental characteristics of monochorionic diamniotic twin pairs with early- and late-onset discordant growth. *American Journal of Obstetrics and Gynecology, 199*(5), doi:10.1016/j.ajog.2008.04.022.

Li, S.-W., Ling, Y., Jin, S., Lin, Y.-F., Chen, Z.-J., Hu, C.-X., et al. (2014). Expression of soluble vascular endothelial growth factor receptor-1 and placental growth factor in fetal growth restriction cases and intervention effect of tetramethylpyrazine. *Asian Pacific journal of tropical medicine, 7*(8), 663-667.

Li, Z., Wu, G., Sher, R. B., Khavandgar, Z., Hermansson, M., Cox, G. A., et al. (2014). Choline kinase beta is required for normal endochondral bone formation. *Biochimica et Biophysica Acta (BBA)-General Subjects, 1840*(7), 2112-2122.

Lindon, J. C., Holmes, E., & Nicholson, J. K. (2006). Metabonomics techniques and applications to pharmaceutical research & development. *Pharmaceutical research, 23*(6), 1075-1088.

Lindon, J. C., Keun, H. C., Ebbels, T. M., Pearce, J. M., Holmes, E., & Nicholson, J. K. (2005). The Consortium for Metabonomic Toxicology (COMET): aims, activities and achievements.

Lindsay, K. L., Hellmuth, C., Uhl, O., Buss, C., Wadhwa, P. D., Koletzko, B., et al. (2015). Longitudinal metabolomic profiling of amino acids and lipids across healthy pregnancy. *PLoS One, 10*(12), e0145794.

Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data* (Vol. 793): John Wiley & Sons.

Lizewska, B., Teul, J., Kuc, P., Lemancewicz, A., Charkiewicz, K., Goscik, J., et al. (2018). Maternal Plasma Metabolomic Profiles in Spontaneous Preterm Birth: Preliminary Results. *Mediators of Inflammation, 2018*, 13, doi:10.1155/2018/9362820.

Lu, W., Bennett, B. D., & Rabinowitz, J. D. (2008). Analytical strategies for LC–MS-based targeted metabolomics. *Journal of Chromatography B, 871*(2), 236-242, doi:https://doi.org/10.1016/j.jchromb.2008.04.031.

Luan, H., Meng, N., Liu, P., Fu, J., Chen, X., Rao, W., et al. (2015). Non-targeted metabolomics and lipidomics LC–MS data from maternal plasma of 180 healthy pregnant women. *GigaScience, 4*(1), 16.

Lucaroni, F., Morciano, L., Rizzo, G., D' Antonio, F., Buonuomo, E., Palombi, L., et al. (2018). Biomarkers for predicting spontaneous preterm birth: an umbrella systematic review. *The Journal of Maternal-Fetal & Neonatal Medicine, 31*(6), 726-734, doi:10.1080/14767058.2017.1297404.

Lundsberg, L. S., Bracken, M. B., & Saftlas, A. F. (1997). Low-to-moderate gestational alcohol use and intrauterine growth retardation, low birthweight, and preterm delivery. *Annals of Epidemiology, 7*(7), 498-508, doi:10.1016/s1047-2797(97)00081-1.

Luthra, G., Vuckovic, I., Bangdiwala, A., Gray, H., Redmon, J., Barrett, E., et al. (2018). First and second trimester urinary metabolic profiles and fetal growth restriction: an exploratory nested case-control study within the infant development and environment study. *BMC Pregnancy Childbirth, 18*(1), 48.

Määttä, J., Sissala, N., Dimova, E. Y., Serpi, R., Moore, L. G., & Koivunen, P. (2018). Hypoxia causes reductions in birth weight by altering maternal glucose and lipid metabolism. *Scientific reports, 8*(1), 13583.

Maehara, K., Kanayama, N., Maradny, E. E., Uezato, T., Fujita, M., & Terao, T. (1996). Mechanical stretching induces interleukin-8 gene expression in fetal membranes: a possible role for the initiation of human parturition. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 70*(2), 191-196.

Magee, L. A., Pels, A., Helewa, M., Rey, E., & von Dadelszen, P. (2014). Diagnosis, evaluation, and management of the hypertensive disorders of pregnancy. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health, 4*(2), 105-145.

Magnusson-Olsson, A., Hamark, B., Ericsson, A., Wennergren, M., Jansson, T., & Powell, T. (2006). Gestational and hormonal regulation of human placental lipoprotein lipase. *Journal of lipid research, 47*(11), 2551-2561.

Mahadevan, S., Shah, S. L., Marrie, T. J., & Slupsky, C. M. (2008). Analysis of metabolomic data using support vector machines. *Analytical Chemistry, 80*(19), 7562-7570.

Maitre, L., Fthenou, E., Athersuch, T., Coen, M., Toledano, M. B., Holmes, E., et al. (2014). Urinary metabolic profiles in early pregnancy are associated with preterm birth and fetal growth restriction in the Rhea mother–child cohort study. *BMC medicine, 12*(1), 110.

Malamitsi-Puchner, A., Briana, D. D., Boutsikou, M., Kouskouni, E., Hassiakos, D., & Gourgiotis, D. (2007). Perinatal circulating visfatin levels in intrauterine growth restriction. *Pediatrics, 119*(6), e1314-e1318.

Many, A., Hill, L. M., Lazebnik, N., & Martin, J. G. (1995). The association between polyhydramnios and preterm delivery. *Obstetrics & Gynecology, 86*(3), 389-391, doi:https://doi.org/10.1016/0029-7844(95)00179-U.

Margolin, A. A., Ong, S.-E., Schenone, M., Gould, R., Schreiber, S. L., Carr, S. A., et al. (2009). Empirical Bayes analysis of quantitative proteomics experiments. *PLoS One, 4*(10), e7454.

Martínez-Domínguez, S. J., Tajada, M., Chedraui, P., & Pérez-López, F. R. (2018). Systematic review and meta-analysis of Spanish studies regarding the association between maternal 25-hydroxyvitamin D levels and perinatal outcomes. *Gynecological Endocrinology, 34*(11), 987-994.

Matarese, G., Moschos, S., & Mantzoros, C. S. (2005). Leptin in immunology. *The Journal of Immunology, 174*(6), 3137-3142.

Matsubara, K., Higaki, T., Matsubara, Y., & Nawa, A. (2015). Nitric oxide and reactive oxygen species in the pathogenesis of preeclampsia. *International journal of molecular sciences, 16*(3), 4600-4614.

Maynard, S. E., Min, J.-Y., Merchan, J., Lim, K.-H., Li, J., Mondal, S., et al. (2003). Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) could contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *Obstetrical & Gynecological Survey, 58*(9), 564-565.

McCarthy, C., & Kenny, L. C. (2016). Therapeutically targeting mitochondrial redox signalling alleviates endothelial dysfunction in preeclampsia. *Scientific reports, 6*, 32683.

McCarthy, F. P., Drewlo, S., Kingdom, J., Johns, E. J., Walsh, S. K., & Kenny, L. C. (2011). Peroxisome proliferator-activated receptor-gamma as a potential therapeutic target in the treatment of preeclampsia. *Hypertension, 58*(2), 280-286, doi:10.1161/HYPERTENSIONAHA.111.172627.

McCarthy, F. P., Ryan, R. M., & Chappell, L. C. (2018). Prospective biomarkers in preterm preeclampsia: A review. *Pregnancy Hypertension, 14*, 72-78, doi:https://doi.org/10.1016/j.preghy.2018.03.010.

McCormick, M. C. (1985). The contribution of low birth weight to infant mortality and childhood morbidity. *New England Journal of Medicine, 312*(2), 82-90.

McCowan, L. M., Figueras, F., & Anderson, N. H. (2018). Evidence-based national guidelines for the management of suspected fetal growth restriction: comparison, consensus, and controversy. *American Journal of Obstetrics and Gynecology, 218*(2, Supplement), S855-S868, doi:https://doi.org/10.1016/j.ajog.2017.12.004.

McCowan, L. M., Harding, J. E., & Stewart, A. W. (2005). Customised birthweight centiles predict SGA pregnancies with perinatal morbidity. *BJOG: An International Journal of Obstetrics & Gynaecology, 112*(8), 1026-1033.

Melo, A., Vieira, C., Barbieri, M., Rosa-e-Silva, A., Silva, A., Cardoso, V., et al. (2010). High prevalence of polycystic ovary syndrome in women born small for gestational age. *Human reproduction, 25*(8), 2124-2131.

Mendez-Figueroa, H., Pedroza, C., Khan, A. M., & Chauhan, S. P. (2016). Small-for-gestational-age infants among uncomplicated pregnancies at term: a secondary analysis of 9 Maternal-Fetal Medicine Units Network studies. *American Journal of Obstetrics and Gynecology, 215*(5), 628. e621-628. e627.

Menon, R., Jones, J., Gunst, P. R., Kacerovsky, M., Fortunato, S. J., Saade, G. R., et al. (2014). Amniotic Fluid Metabolomic Analysis in Spontaneous Preterm Birth. *Reproductive Sciences, 21*(6), 791-803, doi:10.1177/1933719113518987.

Menon, R., Torloni, M. R., Voltolini, C., Torricelli, M., Merialdi, M., Betrán, A. P., et al. (2011). Biomarkers of spontaneous preterm birth: an overview of the literature in the last four decades. *Reproductive Sciences, 18*(11), 1046-1070.

Mesiano, S. (2001). Roles of estrogen and progesterone in human parturition. In *The endocrinology of parturition* (Vol. 27, pp. 86-104): Karger Publishers.

Millar, L. K., Stollberg, J., DeBuque, L., & Bryant-Greenwood, G. (2000). Fetal membrane distention: determination of the intrauterine surface area and distention of the fetal membranes preterm and at term. *American Journal of Obstetrics and Gynecology, 182*(1), 128-134.

Mitsuhashi, S., Ohkuma, A., Talim, B., Karahashi, M., Koumura, T., Aoyama, C., et al. (2011). A congenital muscular dystrophy with mitochondrial structural abnormalities caused by defective de novo phosphatidylcholine biosynthesis. *The American Journal of Human Genetics, 88*(6), 845-851.

Mizugishi, K., Li, C., Olivera, A., Bielawski, J., Bielawska, A., Deng, C.-X., et al. (2007). Maternal disturbance in activated sphingolipid metabolism causes pregnancy loss in mice. *The Journal of clinical investigation, 117*(10), 2993-3006.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*(4), 264-269.

Mol, B. W. J., Roberts, C. T., Thangaratinam, S., Magee, L. A., de Groot, C. J. M., & Hofmeyr, G. J. (2016). Pre-eclampsia. *The Lancet, 387*(10022), 999-1011, doi:https://doi.org/10.1016/S0140-6736(15)00070-7.

Monier, I., Blondel, B., Ego, A., Kaminski, M., Goffinet, F., & Zeitlin, J. (2016). Does the Presence of Risk Factors for Fetal Growth Restriction Increase the Probability of Antenatal Detection? AF rench National Study. *Paediatric and perinatal epidemiology, 30*(1), 46-55.

Moore, T. R., Iams, J. D., Creasy, R. K., Burau, K. D., & Davidson, A. L. (1994). Diurnal and gestational patterns of uterine activity in normal human pregnancy. The Uterine Activity in Pregnancy Working Group. *Obstetrics and gynecology, 83*(4), 517-523.

Moreira, R. S., Magalhães, L. C., & Alves, C. R. L. (2014). Effect of preterm birth on motor development, behavior, and school performance of school-age children: a systematic review. *Jornal de Pediatria (Versão em Português), 90*(2), 119-134, doi:https://doi.org/10.1016/j.jpedp.2013.05.009.

Mori, T. A., Burke, V., Puddey, I. B., Watts, G. F., O'Neal, D. N., Best, J. D., et al. (2000). Purified eicosapentaenoic and docosahexaenoic acids have differential effects on serum lipids and lipoproteins, LDL particle size, glucose, and insulin in mildly hyperlipidemic men. *The American journal of clinical nutrition, 71*(5), 1085-1094.

Morillon, A.-C., Leite, D. F. B., Yakkundi, S., Gethings, L. A., Thomas, G., Baker, P. N., et al. (2021). Glycerophospholipid and detoxification pathways associated with small for gestation age pathophysiology: discovery metabolomics analysis in the SCOPE cohort. *Metabolomics, 17*(1), 5, doi:10.1007/s11306-020-01740-9.

Morillon, A.-C., Williamson, R. D., Baker, P. N., Kell, D. B., Kenny, L. C., English, J. A., et al. (2020a). Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia. *PLoS One, 15*(3), e0230977.

Morillon, A.-C., Yakkundi, S., Thomas, G., Gethings, L. A., Langridge, J. I., Baker, P. N., et al. (2020b). Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomic analysis in the cork pregnancy cohort. [journal article]. *Metabolomics, 16*(2), 19, doi:10.1007/s11306-020-1639-6.

Morillon, A.-C., Yakkundi, S., Thomas, G., Wells, M., Baker, P., & Kenny, L. Untargeted UPLC-MS analysis of potential pesticide and biomarkers of fetal growth restriction. In *14th Annual Conference Metabolomics Society, 2018* (pp. 200)

Mozurkewich, E. L., Luke, B., Avni, M., & Wolf, F. M. (2000). Working conditions and adverse pregnancy outcome: a meta-analysis. *Obstetrics & Gynecology, 95*(4), 623-635.

Mullen, S. F., & Critser, J. K. (2007). The science of cryobiology. In *Oncofertility fertility preservation for cancer survivors* (pp. 83-109): Springer.

Murthi, P. (2014). Review: Placental homeobox genes and their role in regulating human fetal growth. *Placenta, 35*, S46-S50.

Murthi, P., Abumaree, M., & Kalionis, B. (2014). Analysis of homeobox gene action may reveal novel angiogenic pathways in normal placental vasculature and in clinical pregnancy disorders associated with abnormal placental angiogenesis.

Murthi, P., Doherty, V. L., Said, J. M., Donath, S., Brennecke, S. P., & Kalionis, B. (2006). Homeobox gene ESX1L expression is decreased in human pre-term idiopathic fetal growth restriction. *MHR: Basic science of reproductive medicine, 12*(5), 335-340, doi:10.1093/molehr/gal037.

Murthi, P., Kee, M., Gude, N., Brennecke, S., & Kalionis, B. (2005). Fetal growth restriction is associated with increased apoptosis in the chorionic trophoblast cells of human fetal membranes. *Placenta, 26*(4), 329-338.

Murthi, P., Rajaraman, G., Brennecke, S. P., & Kalionis, B. (2011). The role of placental homeobox genes in human fetal growth restriction. *Journal of pregnancy, 2011*.

Myers, J. E., Tuytten, R., Thomas, G., Laroy, W., Kas, K., Vanpoucke, G., et al. (2013). Integrated Proteomics Pipeline Yields Novel Biomarkers for Predicting Preeclampsia. *Hypertension, 61*(6), 1281-1288, doi:doi:10.1161/HYPERTENSIONAHA.113.01168.

Nageotte, M. P., Dorchester, W., Porto, M., Keegan Jr, K. A., & Freeman, R. K. (1988). Quantitation of uterine activity preceding preterm, term, and postterm labor. *American Journal of Obstetrics and Gynecology, 158*(6), 1254-1259.

Nasioudis, D., Doulaveris, G., & Kanninen, T. (2019). Dyslipidemia in pregnancy and maternal-fetal outcome. *Minerva ginecologica, 71*(2), 155-162.

National Institute for Health and Clinical Excellence (NICE) (2011). Hypertension in pregnancy. *NICE clinical guideline 107. London:*

RCOG Press.

Navaratnam, K., Alfirevic, Z., Baker, P. N., Gluud, C., Grüttner, B., Kublickiene, K., et al. (2013). A multi-centre phase IIa clinical study of predictive testing for preeclampsia: improved pregnancy outcomes via early detection (IMPROvED). *BMC Pregnancy Childbirth, 13*(1), 226, doi:10.1186/1471-2393-13-226.

Nemeth, E., Millar, L. K., & Bryant-Greenwood, G. (2000a). Fetal membrane distention: II. Differentially expressed genes regulated by acute distention in vitro. *American Journal of Obstetrics and Gynecology, 182*(1), 60-67.

Nemeth, E., Tashima, L. S., Yu, Z., & Bryant-Greenwood, G. D. (2000b). Fetal membrane distention: I. Differentially expressed genes regulated by acute distention in amniotic epithelial (WISH) cells. *American Journal of Obstetrics and Gynecology, 182*(1), 50-59.

Nevo, O., Soleymanlou, N., Wu, Y., Xu, J., Kingdom, J., Many, A., et al. (2006). Increased expression of sFlt-1 in in vivo and in vitro models of human placental hypoxia is mediated by HIF-1. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 291*(4), R1085-R1093.

Nezar, M. A.-S., El-Baky, A. M. A., Soliman, O. A.-S., Abdel-Hady, H. A.-S., Hammad, A. M., & Al-Haggar, M. S. (2009). Endothelin-1 and leptin as markers of intrauterine growth restriction. *Indian journal of pediatrics, 76*(5), 485-488.

Nomura, R. M. Y., Paiva, L. V., Costa, V. N., Liao, A. W., & Zugaib, M. (2012). Influence of maternal nutritional status, weight gain and energy intake on fetal growth in high-risk pregnancies. *Revista Brasileira de Ginecologia e Obstetrícia, 34*(3), 107-112.

North, R. A., McCowan, L. M., Dekker, G. A., Poston, L., Chan, E. H., Stewart, A. W., et al. (2011). Clinical risk prediction for pre-eclampsia in nulliparous women: development of model in international prospective cohort. *BMJ, 342*, d1875.

Ong, Y. L., Quah, P. L., Tint, M. T., Aris, I. M., Chen, L. W., van Dam, R. M., et al. (2016). The association of maternal vitamin D status with infant birth outcomes, postnatal growth and adiposity in the first 2 years of life in a multi-ethnic Asian population: the Growing Up in Singapore Towards healthy Outcomes (GUSTO) cohort study. *British Journal of Nutrition, 116*(4), 621-631.

Orczyk-Pawilowicz, M., Jawien, E., Deja, S., Hirnle, L., Zabek, A., & Mlynarz, P. (2016). Metabolomics of human amniotic fluid and maternal plasma during normal pregnancy. *PLoS One, 11*(4), e0152740.

Oskovi Kaplan, Z. A., & Ozgu-Erdinc, A. S. (2018). Prediction of Preterm Birth: Maternal Characteristics, Ultrasound Markers, and Biomarkers: An Updated Overview. *Journal of pregnancy, 2018*.

Ou, C.-W., Chen, Z.-Q., Qi, S., & Lye, S. J. (1998). Increased expression of the rat myometrial oxytocin receptor messenger ribonucleic acid during labor requires both mechanical and hormonal signals. *Biology of reproduction, 59*(5), 1055-1061.

Ou, C.-W., Orsino, A., & Lye, S. J. (1997). Expression of connexin-43 and connexin-26 in the rat myometrium during pregnancy and labor is differentially regulated by mechanical and hormonal signals. *Endocrinology, 138*(12), 5398-5407.

Padmanabhan, V., Cardoso, R. C., & Puttabyatappa, M. (2016). Developmental programming, a pathway to disease. *Endocrinology, 157*(4), 1328-1340.

Papageorghiou, A. T., Ohuma, E. O., Altman, D. G., Todros, T., Ismail, L. C., Lambert, A., et al. (2014). International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. *The Lancet, 384*(9946), 869-879, doi:https://doi.org/10.1016/S0140-6736(14)61490-2.

Park, H., Wood, M. R., Malysheva, O. V., Jones, S., Mehta, S., Brannon, P. M., et al. (2017). Placental vitamin D metabolism and its associations with circulating vitamin D metabolites in pregnant women. *The American journal of clinical nutrition, 106*(6), 1439-1448.

Parry, S., Sciscione, A., Haas, D. M., Grobman, W. A., Iams, J. D., Mercer, B. M., et al. (2017). Role of early second-trimester uterine artery Doppler screening to predict small-for-gestational-age babies in nulliparous women. *American Journal of Obstetrics and Gynecology, 217*(5), 594. e591-594. e510.

Patrick, R. P. (2019). Role of phosphatidylcholine-DHA in preventing APOE4-associated Alzheimer's disease. *The FASEB Journal, 33*(2), 1554-1564, doi:https://doi.org/10.1096/fj.201801412R.

Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. [Perspective]. *Nature Reviews Molecular Cell Biology, 13*, 263, doi:10.1038/nrm3314.

Pay, A., Frøen, J., Staff, A., Jacobsson, B., & Gjessing, H. (2016). Prediction of small-for-gestational-age status by symphysis–fundus height: a registry-based population cohort study. *BJOG: An International Journal of Obstetrics & Gynaecology, 123*(7), 1167-1173.

Payne, F., Lim, K., Girousse, A., Brown, R. J., Kory, N., Robbins, A., et al. (2014). Mutations disrupting the Kennedy phosphatidylcholine pathway in humans with congenital lipodystrophy and fatty liver disease. *Proceedings of the National Academy of Sciences, 111*(24), 8901-8906.

Petraglia, F., Giardino, L., Coukos, G., Calza, L., Vale, W., & Genazzani, A. R. (1990). Corticotropin-releasing factor and parturition: plasma and amniotic fluid levels and placental binding sites. *Obstetrics and gynecology, 75*(5), 784-789.

Petrou, S., Abangma, G., Johnson, S., Wolke, D., & Marlow, N. (2009). Costs and Health Utilities Associated with Extremely Preterm Birth: Evidence from the EPICure Study. *Value in Health, 12*(8), 1124-1134, doi:https://doi.org/10.1111/j.1524-4733.2009.00580.x.

Picca, A., Coelho-Junior, H. J., Cesari, M., Marini, F., Miccheli, A., Gervasoni, J., et al. (2019). The metabolomics side of frailty: Toward personalized medicine for the aged. *Experimental gerontology*, 110692.

Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., et al. (2019). Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites, 9*(4), 76.

Platt, M. J. (2014). Outcomes in preterm infants. *Public Health, 128*(5), 399-403, doi:https://doi.org/10.1016/j.puhe.2014.03.010.

Postle, A. D., Al, M. D. M., Burdge, G. C., & Hornstra, G. (1995). The composition of individual molecular species of plasma phosphatidylcholine in human pregnancy. *Early Human Development, 43*(1), 47-58, doi:https://doi.org/10.1016/0378-3782(95)01663-N.

Priyadarshi, A., Verge, C. F., Vandervliet, L., Mackay, D. J. G., & Bolisetty, S. (2015). Transient Neonatal Diabetes Mellitus followed by recurrent asymptomatic hypoglycaemia: a case report. *BMC Pediatrics, 15*, doi:10.1186/s12887-015-0512-7.

Purisch, S. E., DeFranco, E. A., Muglia, L. J., Odibo, A. O., & Stamilio, D. M. (2008). Preterm birth in pregnancies complicated by major congenital malformations: a population-based study. *American Journal of Obstetrics and Gynecology, 199*(3), 287.e281-287.e288, doi:https://doi.org/10.1016/j.ajog.2008.06.089.

Qin, J.-B., Sheng, X.-Q., Wang, H., Chen, G.-C., Yang, J., Yu, H., et al. (2017a). Worldwide prevalence of adverse pregnancy outcomes associated with in vitro fertilization/intracytoplasmic sperm injection among multiple births: a systematic review and meta-analysis based on cohort studies. [journal article]. *Archives of Gynecology and Obstetrics, 295*(3), 577-597, doi:10.1007/s00404-017-4291-2.

Qin, J.-B., Sheng, X.-Q., Wu, D., Gao, S.-Y., You, Y.-P., Yang, T.-B., et al. (2017b). Worldwide prevalence of adverse pregnancy outcomes among singleton pregnancies after in vitro fertilization/intracytoplasmic sperm injection: a systematic review and meta-analysis. [journal article]. *Archives of Gynecology and Obstetrics, 295*(2), 285-301, doi:10.1007/s00404-016-4250-3.

Qin, J., Yang, T., Xiao, S., Tan, H., Feng, T., & Fu, H. (2014). Reported Estimates of Adverse Pregnancy Outcomes among Women with and without Syphilis: A Systematic Review and Meta-Analysis. *PLoS One, 9*(7), e102203, doi:10.1371/journal.pone.0102203.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *URL http://www. R-project. org/.(3.3. 1) Software Vienna, Austria: R Foundation for Statistical Computing*.

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*, 45, doi:10.1038/83496.

Rab, A., Szentpéteri, I., Kornya, L., Börzsönyi, B., Demendi, C., & Joó, J. G. (2013). Placental gene expression patterns of epidermal growth factor in intrauterine growth restriction. *European Journal of Obstetrics & Gynecology and Reproductive Biology, 170*(1), 96-99, doi:https://doi.org/10.1016/j.ejogrb.2013.05.020.

Rajewski, M., & Skrzypczak, J. (2006). Frequency of antiphospholipid antibodies and factor V (G1691A), prothrombin (G20210A) gene polimorphism among women with pregnancy complications. *Polskie Archiwum Medycyny Wewnetrznej, 115*(5), 417-425.

Rambjør, G. S., Windsor, S. L., & Harris, W. S. (1996). Eicosapentaenoic acid is primarily responsible for hypotriglyceridemic effect of fish oil in humans. *Lipids, 31*(1), S45-S49.

Rana, S., Karumanchi, S. A., & Lindheimer, M. D. (2014). Angiogenic factors in diagnosis, management, and research in preeclampsia. *Hypertension, 63*(2), 198-202.

Ravelli, A. C., van der Meulen, J. H., Michels, R., Osmond, C., Barker, D. J., Hales, C., et al. (1998). Glucose tolerance in adults after prenatal exposure to famine. *The Lancet, 351*(9097), 173-177.

Ray, J. G., Jiang, D., Sgro, M., Shah, R., Singh, G., & Mamdani, M. M. (2009). Thresholds for small for gestational age among newborns of East Asian and South Asian ancestry. *Journal of Obstetrics and Gynaecology Canada, 31*(4), 322-330.

Redman, C. W. G. (2014). The six stages of pre-eclampsia. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health, 4*(3), 246, doi:https://doi.org/10.1016/j.preghy.2014.04.020.

Redman, C. W. G., & Sargent, I. L. (2009). Placental Stress and Pre-eclampsia: A Revised View. *Placenta, 30*, 38-42, doi:https://doi.org/10.1016/j.placenta.2008.11.021.

Redman, C. W. G., & Staff, A. C. (2015). Preeclampsia, biomarkers, syncytiotrophoblast stress, and placental capacity. *American Journal of Obstetrics and Gynecology, 213*(4, Supplement), S9.e1-S9.e4, doi:https://doi.org/10.1016/j.ajog.2015.08.003.

Reece, M. S., McGregor, J. A., Allen, K. G. D., & Harris, M. A. (1997). Maternal and perinatal long-chain fatty acids: Possible roles in preterm birth. *American Journal of Obstetrics and Gynecology, 176*(4), 907-914, doi:https://doi.org/10.1016/S0002-9378(97)70620-3.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research, 43*(7), e47-e47.

Roberts, J. M., & Hubel, C. A. (2009). The two stage model of preeclampsia: variations on the theme. *Placenta, 30*, 32-37.

Robinson, W. P., Penaherrera, M. S., Jiang, R., Avila, L., Sloan, J., McFadden, D. E., et al. (2010). Assessing the role of placental trisomy in preeclampsia and intrauterine growth restriction. *Prenatal Diagnosis, 30*(1), 1-8, doi:10.1002/pd.2409.

Rolle-Kampczyk, U. E., Krumsiek, J., Otto, W., Röder, S. W., Kohajda, T., Borte, M., et al. (2016). Metabolomics reveals effects of maternal smoking on endogenous metabolites from lipid metabolism in cord blood of newborns. *Metabolomics, 12*(4), 76.

Roma, E., Arnau, A., Berdala, R., Bergos, C., Montesinos, J., & Figueras, F. (2015). Ultrasound screening for fetal growth restriction at 36 vs 32 weeks' gestation: a randomized trial (ROUTE). *Ultrasound in Obstetrics & Gynecology, 46*(4), 391-397.

Romero, R., Espinoza, J., Chaiworapongsa, T., & Kalache, K. Infection and prematurity and the role of preventive strategies. In *Seminars in Neonatology, 2002* (Vol. 7, pp. 259-274, Vol. 4): Elsevier

Romero, R., Espinoza, J., Gotsch, F., Kusanovic, J., Friel, L., Erez, O., et al. (2006). The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology, 113*(s3), 118-135, doi:doi:10.1111/j.1471-0528.2006.01150.x.

Romero, R., Espinoza, J., Kusanovic, J. P., Gotsch, F., Hassan, S., Erez, O., et al. (2006). The preterm parturition syndrome. *BJOG: An International Journal of Obstetrics & Gynaecology, 113*, 17-42.

Romero, R., Gómez, R., Chaiworapongsa, T., Conoscenti, G., Cheol Kim, J., & Mee Kim, Y. (2001). The role of infection in preterm labour and delivery. *Paediatric and perinatal epidemiology, 15*, 41-56.

Romero, R., Gomez, R., Ghezzi, F., Yoon, B. H., Mazor, M., Edwin, S. S., et al. (1998). A fetal systemic inflammatory response is followed by the spontaneous onset of preterm parturition. *American Journal of Obstetrics and Gynecology, 179*(1), 186-193.

Romero, R., Gonzalez, R., Sepulveda, W., Brandt, F., Ramirez, M., Sorokin, Y., et al. (1992a). Infection and labor: VIII. Microbial invasion of the amniotic cavity in patients with suspected cervical incompetence: prevalence and clinical significance. *American Journal of Obstetrics and Gynecology, 167*(4), 1086-1091.

Romero, R., Mazaki-Tovi, S., Vaisbuch, E., Kusanovic, J. P., Chaiworapongsa, T., Gomez, R., et al. (2010). Metabolomics in premature labor: a novel approach to identify patients at risk for preterm delivery. *The Journal of Maternal-Fetal & Neonatal Medicine, 23*(12), 1344-1359.

Romero, R., Mazor, M., Gomez, R., Gonzalez, R., Galasso, M., & Cotton, D. (1993a). Cervix, incompetence and premature labor. *The fetus, 3*(1), 1-10.

Romero, R., Salafia, C. M., Athanassiadis, A. P., Hanaoka, S., Mazor, M., Sepulveda, W., et al. (1992b). The relationship between acute inflammatory lesions of the preterm placenta and amniotic fluid microbiology. *American Journal of Obstetrics and Gynecology, 166*(5), 1382-1388.

Romero, R., Sepulveda, W., Baumann, P., Yoon, B., Brandt, F., Gomez, R., et al. (1993b). The preterm labor syndrome: biochemical, cytologic, immunologic, pathologic, microbiologic, and clinical evidence that preterm labor is a heterogeneous disease. *Am J Obstet Gynecol, 168*(1 Pt 2), 287.

Romero, R., Yoon, B. H., Kenney, J. S., Gomez, R., Allison, A. C., & Sehgal, P. B. (1993). Amniotic fluid interleukin-6 determinations are of diagnostic and prognostic value in preterm labor. *American journal of reproductive immunology, 30*(2-3), 167-183.

Roos, N., Kieler, H., Sahlin, L., Ekman-Ordeberg, G., Falconer, H., & Stephansson, O. (2011). Risk of adverse pregnancy outcomes in women with polycystic ovary syndrome: population based cohort study. *BMJ, 343*, d6309.

Rosenberg, A. (2008). The IUGR Newborn. *Seminars in Perinatology, 32*(3), 219-224, doi:https://doi.org/10.1053/j.semperi.2007.11.003.

Rutter, C. M., & Gatsonis, C. A. (2001). A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in medicine, 20*(19), 2865-2884.

Sabatine, M. S., Liu, E., Morrow, D. A., Heller, E., McCarroll, R., Wiegand, R., et al. (2005). Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation, 112*(25), 3868-3875.

Saito, S., Shiozaki, A., Nakashima, A., Sakai, M., & Sasaki, Y. (2007). The role of the immune system in preeclampsia. *Molecular aspects of medicine, 28*(2), 192-209.

Sakrak, O., Kerem, M., Bedirli, A., Pasaoglu, H., Akyurek, N., Ofluoglu, E., et al. (2008). Ergothioneine modulates proinflammatory cytokines and heat shock protein 70 in mesenteric ischemia and reperfusion injury. *J Surg Res, 144*(1), 36-42, doi:10.1016/j.jss.2007.04.020.

Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R., & Dunn, W. B. (2013). The role of reporting standards for metabolite annotation and identification in metabolomic studies. *GigaScience, 2*(1), 13-13, doi:10.1186/2047-217X-2-13.

Sánchez-Aranguren, L. C., Prada, C. E., Riaño-Medina, C. E., & Lopez, M. (2014). Endothelial dysfunction and preeclampsia: role of oxidative stress. *Frontiers in physiology, 5*, 372.

Sandman, C. A., Glynn, L., Schetter, C. D., Wadhwa, P., Garite, T., Chicz-DeMet, A., et al. (2006). Elevated maternal cortisol early in pregnancy predicts third trimester levels of placental corticotropin releasing hormone (CRH): priming the placental clock. *peptides, 27*(6), 1457-1463.

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.

Sarafian, M. H., Gaudin, M., Lewis, M. R., Martin, F.-P., Holmes, E., Nicholson, J. K., et al. (2014). Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra

performance liquid Chromatography–Mass spectrometry. *Analytical Chemistry, 86*(12), 5766-5774.

Sattar, N., Greer, I. A., Louden, J., Lindsay, G., McConnell, M., Shepherd, J., et al. (1997). Lipoprotein subfraction changes in normal pregnancy: threshold effect of plasma triglyceride on appearance of small, dense low density lipoprotein. *The Journal of Clinical Endocrinology & Metabolism, 82*(8), 2483-2491.

Sawilowsky, S. S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods, 8*(2), 26.

Schloerke, B., Crowley, J., Cook, D., Hofmann, H., Wickham, H., Briatte, F., et al. (2011). Ggally: Extension to ggplot2.

Schomig, E. (2005). Discovery of the ergothioneine transporter. *Proc Natl Acad Sci USA, 102*, 5256-5261.

Schönfeld, P., Więckowski, M. R., Lebiedzińska, M., & Wojtczak, L. (2010). Mitochondrial fatty acid oxidation and oxidative stress: Lack of reverse electron transfer-associated production of reactive oxygen species. *Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1797*(6), 929-938, doi:https://doi.org/10.1016/j.bbabio.2010.01.010.

Schwarz, H. P., Dahlke, M., & Dreisbach, L. (1977). Phospholipid composition of blood plasma, erythrocytes, and" ghosts" in sickle cell disease. *Clinical Chemistry, 23*(9), 1548-1550.

Sedeek, M., Gilbert, J. S., LaMarca, B. B., Sholook, M., Chandler, D. L., Wang, Y., et al. (2008). Role of Reactive Oxygen Species in Hypertension Produced by Reduced Uterine Perfusion in Pregnant Rats. *American Journal of Hypertension, 21*(10), 1152-1156, doi:10.1038/ajh.2008.239.

Seferovic, M. D., Goodspeed, D. M., Chu, D. M., Krannich, L. A., Gonzalez-Rodriguez, P. J., Cox, J. E., et al. (2015). Heritable IUGR and adult metabolic syndrome are reversible and associated with alterations in the metabolome following dietary supplementation of 1-carbon intermediates. *The FASEB Journal, 29*(6), 2640-2652.

Servillo, L., Castaldo, D., Casale, R., D'Onofrio, N., Giovane, A., Cautela, D., et al. (2015). An uncommon redox behavior sheds light on the cellular antioxidant properties of ergothioneine. *Free Radical Biology and Medicine, 79*, 228-236.

Sharma, D., Farahbakhsh, N., Shastri, S., & Sharma, P. (2016a). Intrauterine growth restriction–part 2. [Review]. *Journal of Maternal-Fetal and Neonatal Medicine, 29*(24), 4037-4048, doi:10.3109/14767058.2016.1154525.

Sharma, D., Shastri, S., Farahbakhsh, N., & Sharma, P. (2016b). Intrauterine growth restriction–part 1. [Review]. *Journal of Maternal-Fetal and Neonatal Medicine, 29*(24), 3977-3987, doi:10.3109/14767058.2016.1152249.

Sharma, S. S., Jangale, N. M., Harsulkar, A. M., Gokhale, M. K., & Joshi, B. N. (2017). Chronic maternal calcium and 25-hydroxyvitamin D deficiency in Wistar rats programs abnormal hepatic gene expression leading to hepatic steatosis in female offspring. *The Journal of nutritional biochemistry, 43*, 36-46.

Shaw, W. (2010). Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of Clostridia spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. *Nutritional neuroscience, 13*(3), 135-143.

Sher, R. B., Aoyama, C., Huebsch, K. A., Ji, S., Kerner, J., Yang, Y., et al. (2006). A rostrocaudal muscular dystrophy caused by a defect in choline kinase beta, the first enzyme in phosphatidylcholine biosynthesis. *Journal of Biological Chemistry, 281*(8), 4938-4948.

Shevchenko, A., & Simons, K. (2010). Lipidomics: coming to grips with lipid diversity. [Perspective]. *Nature Reviews Molecular Cell Biology, 11*, 593, doi:10.1038/nrm2934.

Sideris, I. G., & Nicolaides, K. H. (1990). Amniotic fluid pressure during pregnancy. *Fetal Diagnosis and Therapy, 5*(2), 104-108.

Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G.-Z., et al. (2005). Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical Chemistry, 77*(7), 2187-2200.

Simchen, M. J., Beiner, M. E., Strauss-Liviathan, N., Dulitzky, M., Kuint, J., Mashiach, S., et al. (2000). Neonatal outcome in growth-restricted versus appropriately grown preterm infants. *American Journal of Perinatology, 17*(04), 187-192.

Singer, L. T., Salvator, A., Guo, S., Collin, M., Lilien, L., & Baley, J. (1999). Maternal psychological distress and parenting stress after the birth of a very low-birth-weight infant. [Article]. *Journal of the American Medical Association, 281*(9), 799-805, doi:10.1001/jama.281.9.799.

Sladek, S. M., Westerhausen-Larson, A., & Roberts, J. M. (1999). Endogenous nitric oxide suppresses rat myometrial connexin 43 gap junction protein expression during pregnancy. *Biology of reproduction, 61*(1), 8-13.

Slattery, M. M., & Morrison, J. J. (2002). Preterm delivery. *The Lancet, 360*(9344), 1489-1497, doi:https://doi.org/10.1016/S0140-6736(02)11476-0.

Smith, E., Ottosson, F., Hellstrand, S., Ericson, U., Orho-Melander, M., Fernandez, C., et al. (2019). Ergothioneine is associated with reduced mortality and decreased risk of cardiovascular disease. *Heart*.

Smith, G. C. S., & Pell, J. P. (2001). Teenage pregnancy and risk of adverse perinatal outcomes associated with first and second births: Population based retrospective cohort study. [Article]. *British Medical Journal, 323*(7311), 476-479.

Smith, L. K., Draper, E. S., Manktelow, B. N., Dorling, J. S., & Field, D. J. (2007). Socioeconomic inequalities in very preterm birth rates. *Archives of Disease in Childhood - Fetal and Neonatal Edition, 92*(1), F11-F14, doi:10.1136/adc.2005.090308.

Smits, L. J., & Essed, G. G. (2001). Short interpregnancy intervals and unfavourable pregnancy outcome: role of folate depletion. *The Lancet, 358*(9298), 2074-2077.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology, 3*(1), 1-25.

Snijders, R. J. M., Sherrod, C., Gosden, C. M., & Nicolaides, K. H. (1993). FETAL GROWTH-RETARDATION - ASSOCIATED MALFORMATIONS AND CHROMOSOMAL-ABNORMALITIES. *American Journal of Obstetrics and Gynecology, 168*(2), 547-555.

Soong, Y. K., Wang, T. H., Lee, Y. S., Chen, C. P., Chang, C. L., Ho, S. Y., et al. (2009). GENOME-WIDE DETECTION OF UNIPARENTAL DISOMY IN A FETUS WITH INTRAUTERINE GROWTH RESTRICTION USING GENOTYPING MICROARRAYS. *Taiwanese Journal of Obstetrics & Gynecology, 48*(2), 152-158.

Souza, R. T., Mayrink, J., Leite, D. F., Costa, M. L., Calderon, I. M., Rocha Filho, E. A., et al. (2019). Metabolomics applied to maternal and perinatal health: a review of new frontiers with a translation potential. *Clinics, 74*.

Spicer, R. A., Salek, R., & Steinbeck, C. (2017). Compliance with minimum information guidelines in public metabolomics repositories. [Analysis]. *Scientific Data, 4*, 170137, doi:10.1038/sdata.2017.137

https://www.nature.com/articles/sdata2017137#supplementary-information.

Staneva, A., Bogossian, F., Pritchard, M., & Wittkowski, A. (2015). The effects of maternal depression, anxiety, and perceived stress during pregnancy on preterm birth: A systematic review. *Women and Birth, 28*(3), 179-193, doi:https://doi.org/10.1016/j.wombi.2015.02.003.

Stipoljev, F., Latin, V., Kos, M., Miskovic, B., & Kurjak, A. (2001). Correlation of confined placental mosaicism with fetal intrauterine growth retardation - A case control study

of placentas at delivery. *Fetal Diagnosis and Therapy, 16*(1), 4-9, doi:10.1159/000053871.

Stjernholm, Y., Sahlin, L., Åkerbergb, S., Elinder, A., Eriksson, H. A., Malmström, A., et al. (1996). Cervical ripening in humans: potential roles of estrogen, progesterone, and insulin-like growth factor-I. *American Journal of Obstetrics and Gynecology, 174*(3), 1065-1071.

Stone, W. L., Bailey, B., & Khraisha, N. (2014). The pathophysiology of smoking during pregnancy: a systems biology approach. *Front Biosci (Elite Ed), 6*(2), 318-328.

Strobino, D. M., Ensminger, M. E., Kim, Y. J., & Nanda, J. (1995). MECHANISMS FOR MATERNAL AGE-DIFFERENCES IN BIRTH-WEIGHT. *American Journal of Epidemiology, 142*(5), 504-514.

Sugawara, J., Ochi, D., Yamashita, R., Yamauchi, T., Saigusa, D., Wagata, M., et al. (2019). Maternity Log study: a longitudinal lifelog monitoring and multiomics analysis for the early prediction of complicated pregnancy. *BMJ Open, 9*(2), e025939-e025939, doi:10.1136/bmjopen-2018-025939.

Sulek, K., Han, T.-L., Villas-Boas, S. G., Wishart, D. S., Soh, S.-E., Kwek, K., et al. (2014). Hair metabolomics: identification of fetal compromise provides proof of concept for biomarker discovery. *Theranostics, 4*(9), 953.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis. [journal article]. *Metabolomics, 3*(3), 211-221, doi:10.1007/s11306-007-0082-2.

Swartz, M. E. (2005). UPLC™: an introduction and review. *Journal of Liquid Chromatography & Related Technologies, 28*(7-8), 1253-1263.

Tang, Q., Wu, W., Xu, X., Huang, L., Gao, Q., Chen, H., et al. (2013). miR-141 Contributes to Fetal Growth Restriction by Regulating PLAG1 Expression. *PLoS One, 8*(3), e58737, doi:10.1371/journal.pone.0058737.

Tang, R. M. Y., Cheah, I. K.-M., Yew, T. S. K., & Halliwell, B. (2018). Distribution and accumulation of dietary ergothioneine and its metabolites in mouse tissues. *Scientific reports, 8*(1), 1601.

Tang, W. H. W., Wang, Z., Levison, B. S., Koeth, R. A., Britt, E. B., Fu, X., et al. (2013). Intestinal Microbial Metabolism of Phosphatidylcholine and Cardiovascular Risk. *New England Journal of Medicine, 368*(17), 1575-1584, doi:10.1056/NEJMoa1109400.

Tanret C (1909). Sur une base nouvelle retiree du seigle ergote, l'ergothioneine. *Compt Rend*, 222-224.

Thomas, M. M., Haghiac, M., Grozav, C., Minium, J., Calabuig-Navarro, V., & O'Tierney-Ginn, P. (2018). Oxidative Stress Impairs Fatty Acid Oxidation and Mitochondrial Function in the Term Placenta. *Reproductive Sciences*, 1933719118802054.

Thompson, J. M. D., Irgens, L. M., Rasmussen, S., & Daltveit, A. K. (2006). Secular trends in socio-economic status and the implications for preterm birth. *Paediatric and perinatal epidemiology, 20*(3), 182-187, doi:10.1111/j.1365-3016.2006.00711.x.

Tian, Y., Zhou, R., Rehg, J. E., & Jackowski, S. (2007). Role of phosphocholine cytidylyltransferase α in lung development. *Molecular and cellular biology, 27*(3), 975-982.

Torry, D. S., Mukherjea, D., Arroyo, J., & Torry, R. J. (2003). Expression and function of placenta growth factor: implications for abnormal placentation. *Journal of the Society for Gynecologic Investigation, 10*(4), 178-188.

Underwood, B. R., Broadhurst, D., Dunn, W. B., Ellis, D. I., Michell, A. W., Vacher, C., et al. (2006). Huntington disease patients and transgenic mice have similar pro-catabolic serum metabolite profiles. *Brain, 129*(4), 877-886.

Uran, S., Larsen, Å., Jacobsen, P. B., & Skotland, T. (2001). Analysis of phospholipid species in human blood using normal-phase liquid chromatography coupled with electrospray

ionization ion-trap tandem mass spectrometry. *Journal of Chromatography B: Biomedical Sciences and Applications, 758*(2), 265-275, doi:https://doi.org/10.1016/S0378-4347(01)00188-8.

Vadillo, F. O., Avila, M. V., Hernández, C. G., Arechavaleta, F. V., & Beltrán, J. M. (2000). Apoptosis in trophoblast of patients with recurrent spontaneous abortion of unidentified cause. *Ginecologia y obstetricia de Mexico, 68*, 122-131.

Vaka, V. R., McMaster, K. M., Cunningham, M. W., Jr., Ibrahim, T., Hazlewood, R., Usry, N., et al. (2018). Role of Mitochondrial Dysfunction and Reactive Oxygen Species in Mediating Hypertension in the Reduced Uterine Perfusion Pressure Rat Model of Preeclampsia. *Hypertension (Dallas, Tex. : 1979), 72*(3), 703-711, doi:10.1161/HYPERTENSIONAHA.118.11290.

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

van der Veen, J. N., Kennelly, J. P., Wan, S., Vance, J. E., Vance, D. E., & Jacobs, R. L. (2017). The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica et Biophysica Acta (BBA) - Biomembranes, 1859*(9, Part B), 1558-1572, doi:https://doi.org/10.1016/j.bbamem.2017.04.006.

van Dijk, M., & Oudejans, C. B. M. (2011). STOX1: Key Player in Trophoblast Dysfunction Underlying Early Onset Preeclampsia with Growth Retardation. *Journal of pregnancy, 2011*, 7, doi:10.1155/2011/521826.

van Eijsden, M., Hornstra, G., van der Wal, M. F., Vrijkotte, T. G., & Bonsel, G. J. (2008a). Maternal n− 3, n− 6, and trans fatty acid profile early in pregnancy and term birth weight: a prospective cohort study. *The American journal of clinical nutrition, 87*(4), 887-895.

van Eijsden, M., Smits, L. J., van der Wal, M. F., & Bonsel, G. J. (2008b). Association between short interpregnancy intervals and term birth weight: the role of folate depletion. *The American journal of clinical nutrition, 88*(1), 147-153, doi:10.1093/ajcn/88.1.147.

Vavich, J. M., & Howell, R. R. (1970). Ultraviolet-absorbing compounds in urine of normal newborns and young children. *Clinical Chemistry, 16*(8), 702-706.

Verlijsdonk, J.-W., Winkens, B., Boers, K., Scherjon, S., & Roumen, F. (2012). Suspected versus non-suspected small-for-gestational age fetuses at term: perinatal outcomes. *The Journal of Maternal-Fetal & Neonatal Medicine, 25*(7), 938-943.

Vesterinen, H. M., Morello-Frosch, R., Sen, S., Zeise, L., & Woodruff, T. J. (2017). Cumulative effects of prenatal-exposure to exogenous chemicals and psychosocial stress on fetal growth: systematic-review of the human and animal evidence. *PLoS One, 12*(7), e0176331.

Villas-Boas, S. G., Nielsen, J., Smedsgaard, J., Hansen, M. A., & Roessner-Tunali, U. (2007). *Metabolome analysis: an introduction* (Vol. 24): John Wiley & Sons.

Visentin, S., Crotti, S., Donazzolo, E., D'Aronco, S., Nitti, D., Cosmi, E., et al. (2017). Medium chain fatty acids in intrauterine growth restricted and small for gestational age pregnancies. *Metabolomics, 13*(5), 54.

von Dadelszen, P., Payne, B., Li, J., Ansermino, J. M., Pipkin, F. B., Côté, A.-M., et al. (2011). Prediction of adverse maternal outcomes in pre-eclampsia: development and validation of the fullPIERS model. *The Lancet, 377*(9761), 219-227.

Vrachnis, N., Kalampokas, E., Sifakis, S., Vitoratos, N., Kalampokas, T., Botsis, D., et al. (2013). Placental growth factor (PlGF): a key to optimizing fetal growth. *The Journal of Maternal-Fetal & Neonatal Medicine, 26*(10), 995-1002.

Wadhwa, P. D., Culhane, J. F., Rauh, V., & Barve, S. S. (2001). Stress and preterm birth: neuroendocrine, immune/inflammatory, and vascular mechanisms. *Maternal and child health journal, 5*(2), 119-125.

Walker, P. G. T., ter Kuile, F. O., Garske, T., Menendez, C., & Ghani, A. C. (2014). Estimated risk of placental infection and low birthweight attributable to Plasmodium falciparum malaria in Africa in 2010: a modelling study. *Lancet Global Health, 2*(8), E460-E467.

Wallenstein, M. B., Harper, L. M., Odibo, A. O., Roehl, K. A., Longman, R. E., Macones, G. A., et al. (2012). Fetal congenital heart disease and intrauterine growth restriction: a retrospective cohort study. *Journal of Maternal-Fetal & Neonatal Medicine, 25*(6), 662-665, doi:10.3109/14767058.2011.597900.

Walsh, S. K., English, F. A., Johns, E. J., & Kenny, L. C. (2009). Plasma-mediated vascular dysfunction in the reduced uterine perfusion pressure model of preeclampsia: a microvascular characterization. *Hypertension, 54*(2), 345-351.

Wang, F., Mukherjee, S., Richardson, S., & Hill, S. M. (2018). High-dimensional regression in practice: an empirical study of finite-sample prediction, variable selection and ranking. *arXiv preprint arXiv:1808.00723*.

Wang, L., Magdaleno, S., Tabas, I., & Jackowski, S. (2005). Early embryonic lethality in mice with targeted deletion of the CTP: phosphocholine cytidylyltransferase α gene (Pcyt1a). *Molecular and cellular biology, 25*(8), 3357-3363.

Wang, N., Wang, X., Li, Q., Han, B., Chen, Y., Zhu, C., et al. (2017). The famine exposure in early life and metabolic syndrome in adulthood. *Clinical nutrition, 36*(1), 253-259.

Wang, Y., Adgent, M., Su, P.-H., Chen, H.-Y., Chen, P.-C., Hsiung, C. A., et al. (2016). Prenatal exposure to perfluorocarboxylic acids (PFCAs) and fetal and postnatal growth in the Taiwan Maternal and Infant Cohort Study. *Environmental health perspectives, 124*(11), 1794-1800.

Want, E. J., Wilson, I. D., Gika, H., Theodoridis, G., Plumb, R. S., Shockcor, J., et al. (2010). Global metabolic profiling procedures for urine using UPLC–MS. *Nature protocols, 5*(6), 1005-1018.

Watson, A. D. (2006). Lipidomics: a global approach to lipid analysis in biological systems. *Journal of lipid research, 47*(10), 2101.

Watts, D., Krohn, M. A., Hillier, S. L., & Eschenbach, D. A. (1992). The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. *Obstetrics and gynecology, 79*(3), 351-357.

Weiss, S. B., Smith, S. W., & Kennedy, E. P. (1958). The enzymatic formation of lecithin from cytidine diphosphate choline and D-1, 2-diglyceride. *Journal of Biological Chemistry, 231*(1), 53-64.

Wendt, A., Gibbs, C. M., Peters, S., & Hogue, C. J. (2012). Impact of Increasing Inter-pregnancy Interval on Maternal and Infant Health. *Paediatric and perinatal epidemiology, 26*(s1), 239-258, doi:10.1111/j.1365-3016.2012.01285.x.

West, A. A., Yan, J., Jiang, X., Perry, C. A., Innis, S. M., & Caudill, M. A. (2013). Choline intake influences phosphatidylcholine DHA enrichment in nonpregnant women but not in pregnant women in the third trimester. *The American journal of clinical nutrition, 97*(4), 718-727.

Whiting, P. F., Rutjes, A., Westwood, M., Mallet, S., Deeks, J., Reitsma, J., et al. (2011). Research and reporting methods accuracy studies. *Ann Intern Med, 155*(4), 529-536.

WHO (1977). World Health Organization: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. *Acta Obstet Gynecol Scand, 56*, 247-253.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics, 3*(2), 180-185, doi:10.1002/wics.147.

Wilcox, M. A., Smith, S. J., Johnson, I. R., Maynard, P. V., & Chilvers, C. E. D. (1995). The effect of social deprivation on birthweight, excluding physiological and pathological effects. *BJOG: An International Journal of Obstetrics & Gynaecology, 102*(11), 918-924, doi:10.1111/j.1471-0528.1995.tb10882.x.

Wilkins-Haug, L., Quade, B., & Morton, C. C. (2006). Confined placental mosaicism as a risk factor among newborns with fetal growth restriction. *Prenatal Diagnosis, 26*(5), 428-432, doi:10.1002/pd.1430.

Williams, P. J., & Pipkin, F. B. (2011). The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. *Best Practice & Research Clinical Obstetrics & Gynaecology, 25*(4), 405-417.

Williamson Rachel, D., McCarthy Fergus, P., Manna, S., Groarke, E., Kell Douglas, B., Kenny Louise, C., et al. (2020). L-(+)-Ergothioneine Significantly Improves the Clinical Characteristics of Preeclampsia in the Reduced Uterine Perfusion Pressure Rat Model. *Hypertension, 75*(2), 561-568, doi:10.1161/HYPERTENSIONAHA.119.13929.

Williamson, R. D., McCarthy, C., McCarthy, F. P., & Kenny, L. C. (2017). Oxidative stress in pre-eclampsia; have we been looking in the wrong place?

Williamson, R. D., McCarthy, F. P., Khashan, A. S., Totorika, A., Kenny, L. C., & McCarthy, C. (2018). Exploring the role of mitochondrial dysfunction in the pathophysiology of pre-eclampsia. *Pregnancy Hypertension, 13*, 248-253.

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nature reviews Drug discovery, 15*(7), 473.

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic acids research, 46*(D1), D608-D617, doi:10.1093/nar/gkx1089.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58*(2), 109-130, doi:https://doi.org/10.1016/S0169-7439(01)00155-1.

Wood, P. L., & Cebak, J. E. (2018). Lipidomics biomarker studies: errors, limitations, and the future. *Biochemical and biophysical research communications, 504*(3), 569-575.

Wu, G., Aoyama, C., Young, S. G., & Vance, D. E. (2008). Early embryonic lethality caused by disruption of the gene for choline kinase α, the first enzyme in phosphatidylcholine biosynthesis. *Journal of Biological Chemistry, 283*(3), 1456-1462.

Wu, G., Fang, Y.-Z., Yang, S., Lupton, J. R., & Turner, N. D. (2004). Glutathione Metabolism and Its Implications for Health. *The Journal of nutrition, 134*(3), 489-492, doi:10.1093/jn/134.3.489.

Wu, P., Haththotuwa, R., Kwok, C. S., Babu, A., Kotronias, R. A., Rushton, C., et al. (2017). Preeclampsia and Future Cardiovascular Health. *Circulation: Cardiovascular Quality and Outcomes, 10*(2), e003497, doi:doi:10.1161/CIRCOUTCOMES.116.003497.

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics, 9*(2), 280-299.

Xiao, D. H., Li, H. G., & Zang, H. (2015). A Multi-variable Alarm Threshold Optimization Method for Chemical Processes. *Control Engineering of China, 22*(2), 246-251.

Yajnik, C. S., Chandak, G. R., Joglekar, C., Katre, P., Bhat, D. S., Singh, S. N., et al. (2014). Maternal homocysteine in pregnancy and offspring birthweight: epidemiological associations and Mendelian randomization analysis. *International journal of epidemiology, 43*(5), 1487-1497.

Yan, J., Jiang, X., West, A. A., Perry, C. A., Malysheva, O. V., Brenna, J. T., et al. (2013). Pregnancy alters choline dynamics: results of a randomized trial using stable isotope methodology in pregnant and nonpregnant women. *The American journal of clinical nutrition, 98*(6), 1459-1467.

Yang, Q. H., Witkiewicz, B. B., Olney, R. S., Liu, Y. C., Davis, M., Khoury, M. J., et al. (2001). A case-control study of maternal alcohol consumption and intrauterine growth retardation. *Annals of Epidemiology, 11*(7), 497-503, doi:10.1016/s1047-2797(01)00240-x.

Yen, C.-L. E., Stone, S. J., Koliwad, S., Harris, C., & Farese, R. V. (2008). Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *Journal of lipid research, 49*(11), 2283-2301.

Yoon, B. H., Romero, R., Kim, C. J., Jun, J. K., Gomez, R., Choi, J.-H., et al. (1995). Amniotic fluid interleukin-6: a sensitive test for antenatal diagnosis of acute inflammatory lesions of preterm placenta and prediction of perinatal morbidity. *American Journal of Obstetrics and Gynecology, 172*(3), 960-970.

Yoon, B. H., Romero, R., Kim, K. S., Park, J. S., Ki, S. H., Kim, B. I., et al. (1999). A systemic fetal inflammatory response and the development of bronchopulmonary dysplasia. *American Journal of Obstetrics and Gynecology, 181*(4), 773-779.

Yoon, B. H., Romero, R., Park, J. S., Kim, C. J., Kim, S. H., Choi, J.-H., et al. (2000). Fetal exposure to an intra-amniotic inflammation and the development of cerebral palsy at the age of three years. *American Journal of Obstetrics and Gynecology, 182*(3), 675-681.

Yu, K. M. (1992). [Relation between placental morphometry and fetal growth]. *Zhonghua fu chan ke za zhi, 27*(4), 217-219, 250.

Zelena, E., Dunn, W. B., Broadhurst, D., Francis-McIntyre, S., Carroll, K. M., Begley, P., et al. (2009). Development of a robust and repeatable UPLC− MS method for the long-term metabolomic study of human serum. *Analytical Chemistry, 81*(4), 1357-1364.

Zhang, L., Hatzakis, E., & Patterson, A. D. (2016). NMR-Based Metabolomics and Its Application in Drug Metabolism and Cancer Research. [journal article]. *Current Pharmacology Reports, 2*(5), 231-240, doi:10.1007/s40495-016-0067-9.

Zhang, S., Regnault, T. R. H., Barker, P. L., Botting, K. J., McMillen, I. C., McMillan, C. M., et al. (2015). Placental Adaptations in Growth Restriction. *Nutrients, 7*(1), 360-389, doi:10.3390/nu7010360.

Zheng, W., Huang, W., Zhang, L., Tian, Z., Wang, T., Zhang, T., et al. (2018). Changes in serum lipid levels during pregnancy and association with neonatal outcomes: a large cohort study. *Reproductive Sciences, 25*(9), 1406-1412.

Zuckerman, B., Frank, D. A., Hingson, R., Amaro, H., Levenson, S. M., Kayne, H., et al. (1989). EFFECTS OF MATERNAL MARIJUANA AND COCAINE USE ON FETAL GROWTH. *New England Journal of Medicine, 320*(12), 762-768, doi:10.1056/nejm198903233201203.

Appendix A

Dissemination of work

A.1. Peer-reviewed publications

Year	Title
Published manuscripts	
2018	Metabolomics for predicting fetal growth restriction: protocol for a systematic review and meta-analysis. Leite DFB, **Morillon A**, Melo Júnior EF, *et al*. BMJ Open 2018;8:e022743. doi: 10.1136/bmjopen-2018-022743
2019	Examining the predictive accuracy of metabolomics for small-for-gestational-age babies: a systematic review. Leite DFB, **Morillon A**, Melo Júnior EF, *et al*. BMJ Open 2019;9:e031238. doi: 10.1136/bmjopen-2019-031238
2020	Association between phospholipid metabolism in plasma and spontaneous preterm birth: a discovery lipidomic analysis in the cork pregnancy cohort. **Morillon, A.-C**., et al. (2020). Metabolomics 16(2): 19.
2020	Effect of L-Ergothioneine on the metabolic plasma profile of the RUPP rat model of pre-eclampsia. **Morillon, A.-C**, et al. (2020). PLoS One 15(3): e0230977.
2021	Glycerophospholipid and Detoxification Pathways Associated with Small for Gestation Age Pathophysiology: Discovery Metabolomics Analysis in the SCOPE Cohort. **Morillon, A.-C,** et al. (2021). Metabolomics 17(1), 5. doi:10.1007/s11306-020-01740-9
Manuscripts submitted	
2020	Altered plasma levels of triacylglycerol and phosphatidylcholine in preterm birth; a biomarker discovery and validation study in Cork SCOPE birth cohort. K G. Dowling, **A-C Morillon**, et al. Submitted to the American Journal of Obstetrics and Gynecology.

A.2. Conference poster presentations

Year	Title
2019	**Aude-Claire Morillon**, Debora F. B. Leite, Gregoire Thomas, Philip N. Baker, Fergus McCarthy, Louise C. Kenny, Shirish Yakkundi Untargeted UPLC-MS analysis of plasma samples shows higher levels of glycerophospholipids associated with fetal growth restriction. *Irish Mass Spectrometry Society Conference 2019.*
2018	**A-C Morillon**, S Yakkundi, G Thomas, P Baker, L Kenny. Untargeted UPLC-MS analysis of potential pesticide and biomarkers of fetal growth restriction. *Metabolomics Society conference 2018.*
2018	S Yakkundi, L A Gethings, **A-C Morillon**, J Langridge, L Kenny. LIPIDS: Potential biomarkers involved in spontaneous pre-term birth (sp-PTB) – An untargeted metabolomics approach using label free LC-DIA-MS method. *American Society for Mass Spectrometry conference 2018.*
2017	**A-C Morillon**, S Yakkundi, M Wells, P Baker, L Kenny. Pesticide exposure in early pregnancy associated with fetal growth restriction. *Metabolomics Society conference 2017.*
2017	S Yakkundi, L A Gethings, **A-C Morillon**, J Langridge, L Kenny. A label-free LC-DIA-MS method to identify potential biomarkers involved in spontaneous pre-term birth (sp-PTB), using an untargeted metabolomics approach. *American Society for Mass Spectrometry conference 2017.*
2017	L A Gethings, S Yakkundi, G Thomas, **A-C Morillon**, J I Langridge, L Kenny. Identification and label-free quantification of lipid based markers for pre-term birth using a novel scanning quadrupole DIA acquisition method. *American Society for Mass Spectrometry conference 2017.*
2016	S Yakkundi, L A Gethings, **A-C Morillon**, J Langridge, L Kenny. An untargeted metabolomics approach for to identify potential biomarkers in spontaneous pre-term birth (sp-PTB) delivery, using a label-free LC-DIA-MS approach. *American Society for Mass Spectrometry conference 2016; Metabolomics Society conference 2016; INFANT Research Day 2016.*

A.3. Other achievements

Prizes and awards

Year	Title
2018	College of Medicine and Health Travel Bursary of 1000 €

Outreach

Year	Title
2019	Native Scientist – Workshop for children in French
2018	Native Scientist – Workshop for children in French
2017	Two high school visits to promote STEM

Appendix B

Supplementary information for Chapter 2

List of excluded studies and reasons

Authors/ year	Country of enrollment	Additional comments
Exclusions according to study design or statistical analysis		
Barnes CM et al, 2010	United States	Maternal samples collected at delivery.
Bobinski R. 2013	Poland	Cross-sectional study.
Bobinski R. 2014	Poland	Cross-sectional study.
Cao WC et al, 2016	China	Cross-sectional study. The metabolomics technique was not applied.
Chen TT et al, 2017	China	Cross-sectional study.
Cinelli et al, 2018	Italy	
D'Anna R et al, 2004	Italy	Cross-sectional study. The metabolomics technique was not applied.
Guo H et al, 2014	China	Cross-sectional study.
Guo J et al, 2016	China	Cross-sectional study.
Maekawa R et al, 2017	Japan	Cross-sectional study.

Mao D et al, 2010	China	Cross-sectional study.
Miranda J et al 2018	Spain	Cross-sectional study.
Powell et al, 2018	Australia	SGA babies not suspected before birth were considered healthy infants.
Spanou L. et al, 2017	Greece	Cross-sectional study.
Stein TP et al, 2008	United States	Newborns with birth defects were included in the analysis.
Tang R et al, 2013	China	Cross-sectional study.
Visentin S et al, 2017	Italy	Maternal samples collected after clinical recognition of FGR/SGA.
Zhu Y et al, 2018	China	Cross-sectional study.
Zota AR et al, 2009	United States	Cross-sectional study. The metabolomics technique was not applied.

Studies that have not applied metabolomics technique		
Baker PN, 2009	United Kingdom	
Berkowitz GS et al, 2004	United States	

Bodnar LM et al, 2012	United States	
Braun JM et al, 2011	United States	There is no data about FGR.
Cetin I et al, 2002	Italy	
Chong MFF et al, 2015	Singapore	There is no data about birth weight.
Colapinto CK et al, 2015	Canada	The metabolomics technique was not applied for pregnant women's specimens.
Cupul-Uicab LA et al, 2013	United States	
Fruscalzo A et al, 2015	Italy	There is no data about birth weight.
Jusko TA et al, 2006	United States	
Koepke R et al, 2004	Mexico	
Lõpez-Alarcõn M et al, 2015	Mexico	There is no data about birth weight.
Maruta E et al, 2017	Japan	
Miranda ML et al, 2015	United States	
Morley R et al, 2006	Australia	

Muthayya S et al, 2006	India	
Paşaoğlu H et al, 2003	Turkey	
Rahman A et al, 2009	Bangladesh	
Rajasingam D et al, 2009	United Kingdom	
Savitz DA et al, 2002	United States	The metabolomics technique was not applied for pregnant women's specimens.
Savvidou MD et al, 2003	United Kingdom	
Schneuer FJ et al, 2014	Australia	
Snijder CA et al, 2013	Netherlands	
Sweeney AM & Symanski E, 2007	United States	
Takimoto H et al, 2007	Japan	
Terrell ML et al, 2015	United States	
Wei Y et al, 2017	Bangladesh	
Weisskopf MG et al, 2005	United States	

Whyatt RM et al, 2009	United States
Xue F et al, 2007	United States

Studies that have not presented specific data about FGR/SGA	
Bach CC et al, 2016	Denmark
Bachkangi P et al.	United Kingdom
Bahado-Singh RO et al, 2012	United Kingdom
Bahado-Singh RO et al, 2015	United Kingdom
Bahado-Singh RO et al, 2017	United Kingdom
Bentley-Lewis R, 2015	United States
Braun JM et al, 2009	United States
Buckley JP et al, 2016	United States
Cantonwine D et al, 2010	Mexico

Cantonwine D et al, 2015	United States
Casas M et al, 2016	Spain
Castorina R et al, 2017 (a)	United States
Chou WC et al, 2014.	Taiwan
Cunha Figueiredo AC et al, 2017	Brazil
Dalsager L et al, 2018	Denmark
De Renzy-Martin KT. et al, 2014	Poland
Debost-Legrand A et al, 2016	France
Desert et al, 2015	France
Diaz SO et al, 2011	Portugal
Diaz SO et al, 2013	Portugal
Dobierzewska A et al, 2017	Chile
Dudzik D et al, 2015	Spain.

Engström KS et al, 2010	Bangladesh
Ettinger AS et al, 2017	Canada
Feng L et al, 2016	China
Ferguson KK et al, 2014	United States
Ferguson KK et al, 2015	United States
Ferguson KK et al, 2017	United States
Finkelstein JL et al, 2015	United States
Fischer ST et al, 2017	United States
Gao H et al, 2017	China
Gardner RM et al, 2011	Bangladesh
Ghartey J et al, 2017	United States
Graça G et al, 2010	Portugal
Graça G et al, 2012	Portugal

Graça G et al, 2012 (b)	Portugal
Hogeveen M et al, 2010	Netherlands
Huang J et al, 2017	China
Kalhan SC et al, 2003	United States
Khalil AA et al, 2013	United Kingdom
Kuc S et al, 2014	Netherlands
Lenters V et al, 2013	Greenland, Poland, Ukraine
Lenters V et al, 2016	Greenland, Poland, Ukraine
Liu K et al, 2017	China
Lopez-Espinosa MJ et al, 2015	Spain
Marchlewicz EH et al, 2016	United States
Minatoya M et al, 2017	Japan
Minatoya M et al, 2017 (b)	Japan

Minatoya M et al, 2018	Japan	
Murphy MM et al, 2007	Spain	There is no data about any pregnancy outcomes.
Odibo AO et al, 2011	United States	
Pinney SE et al, 2017	United States	
Polanska K et al, 2014	Poland	
Polanska K et al, 2014 (b)	Poland	
Porter A et al, 2018	United States	
Rejc B et al, 2016	Slovenia	
Rijvers CAH et al, 2013	Netherlands	
Robledo C et al, 2013	United States	
Sachse D et al, 2012	Norway	
Scholtens DM et al, 2016	United Kingdom	
Shisler S et al, 2017	United States	Not all analysis were performed with metabolomics approach.

Tamblyn JA et al, 2018	Ireland	Duplicate data. Check Kiely ME et al, 2016.
Thomas MM et al, 2015	New Zealand	
Van Lee L et al, 2015	Singapore	
Virgiliou C et al, 2017	Greece	
Walsh J et al, 2012	Ireland	
Wang PW et al, 2015	Taiwan	
Watkins DJ et al, 2016	United States	
Wolff MS et al, 2008	United States	
Woods MM et al, 2017	United States	
Yang P et al, 2018	China	

Duplicate data		
Horgan R et al, 2009	Australia	Check Horgan R et al, 2011.

Horgan R et al, 2011	Australia	Check Horgan R et al, 2011.
Khashan AS et al, 2013	Ireland	Check Kiely ME et al, 2016.
Sulek et al, 2014	Singapore	Check Sulek et al, 2014.

Individual QUADAS-2 data for all 15 included studies

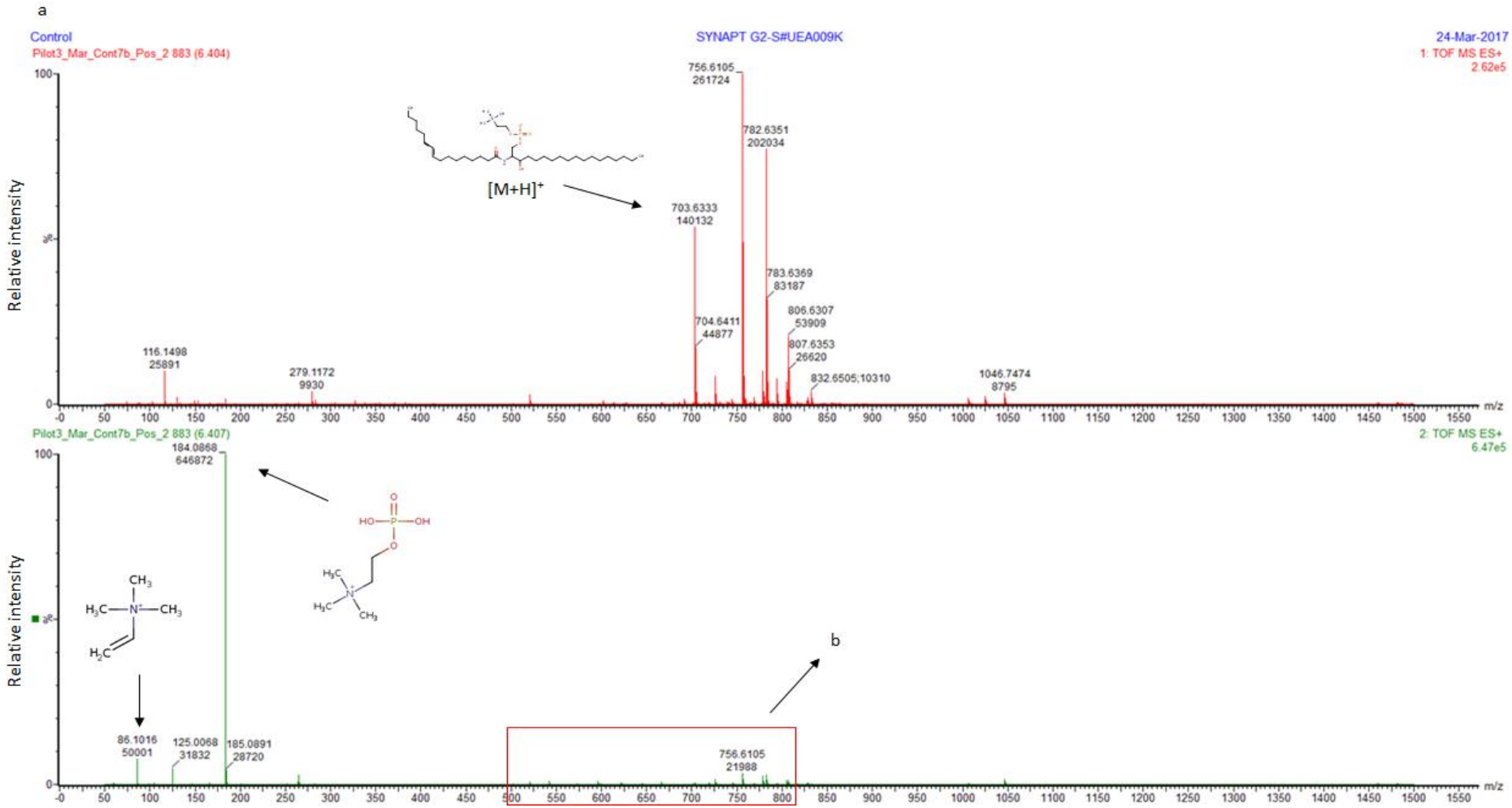
Studies	Risk of bias							
	Patient selection		Index test		Reference standard		Flow and timing	
	Was a consecutive or random sample of patients enrolled?	Did the study avoid inappropriate exclusions?	Were the index test results interpreted without knowledge of the results of the reference standard?	If a threshold was used, was it pre-specified?	Is the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Did all patients receive the same reference standard?	Were all patients included in the analysis?
Grandone E et al, 2006	Yes	Yes	Unclear	No	No	Unclear	Yes	Yes
van Eijsden M et al, 2008	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Horgan R et al, 2011	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Costet N et al, 2012	Yes	Yes	Yes	No	Yes	Unclear	Yes	No
Ertl R et al, 2012	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Gernand AD et al, 2013	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes

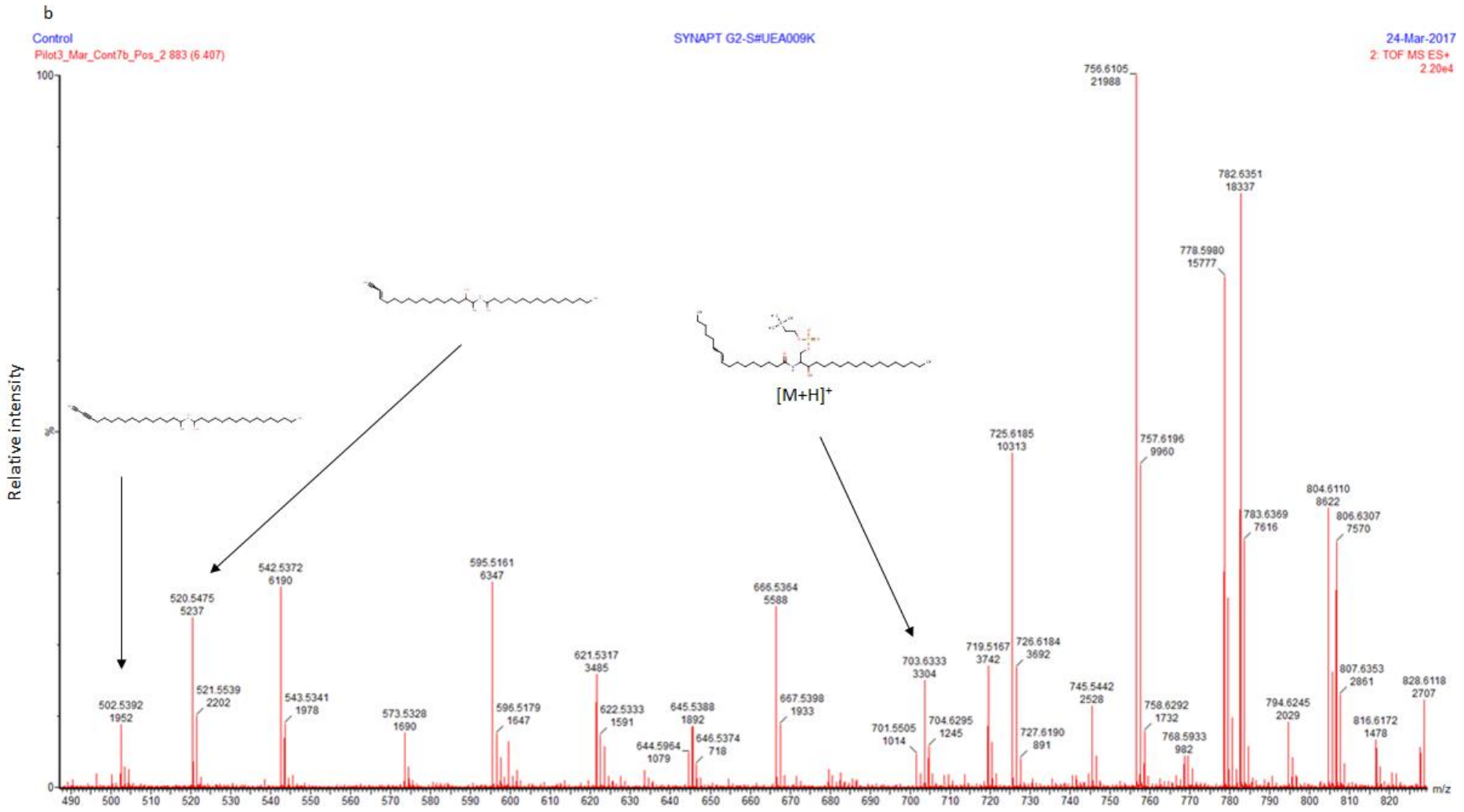
Sulek K et al, 2014	Yes	Yes	No	Yes	Unclear	Unclear	Yes	Yes
Choi R et al, 2016	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes
Kiely ME et al, 2016	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Ong YL et al, 2016	Yes	Yes	No	Yes	Yes	Yes	Yes	No
Wang Y et al, 2016	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Delplancke TDJ et al, 2018	Yes	Yes	No	Yes	Yes	Unclear	Yes	Yes
Luthra G et al, 2018	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Gong S et al, 2018	No	Yes	No	No	Yes	Unclear	Yes	Yes
Morillon AC et al, 2018	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes

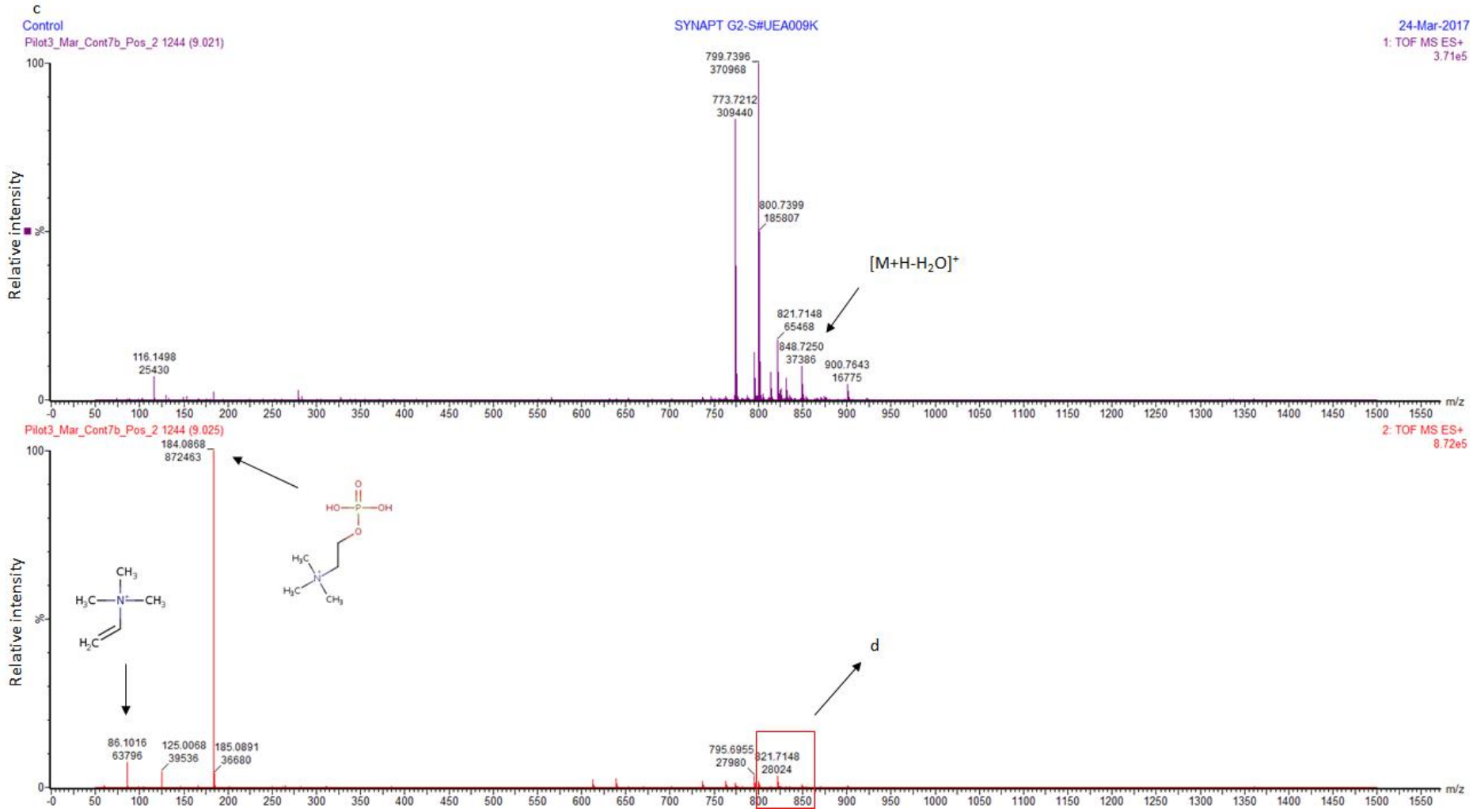
Appendix C

Supplementary information for Chapter 4

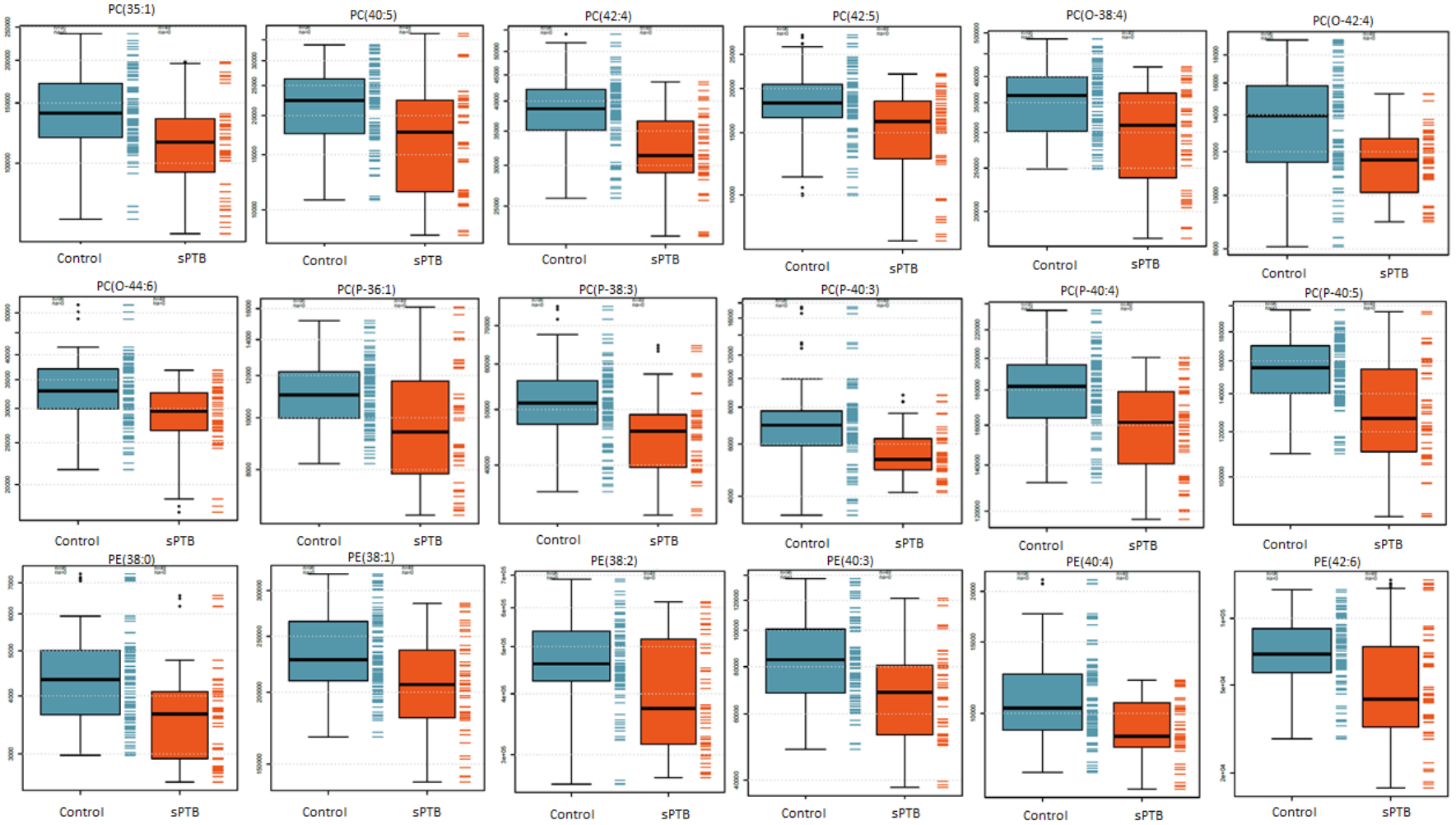
Example of manual validation of two lipids

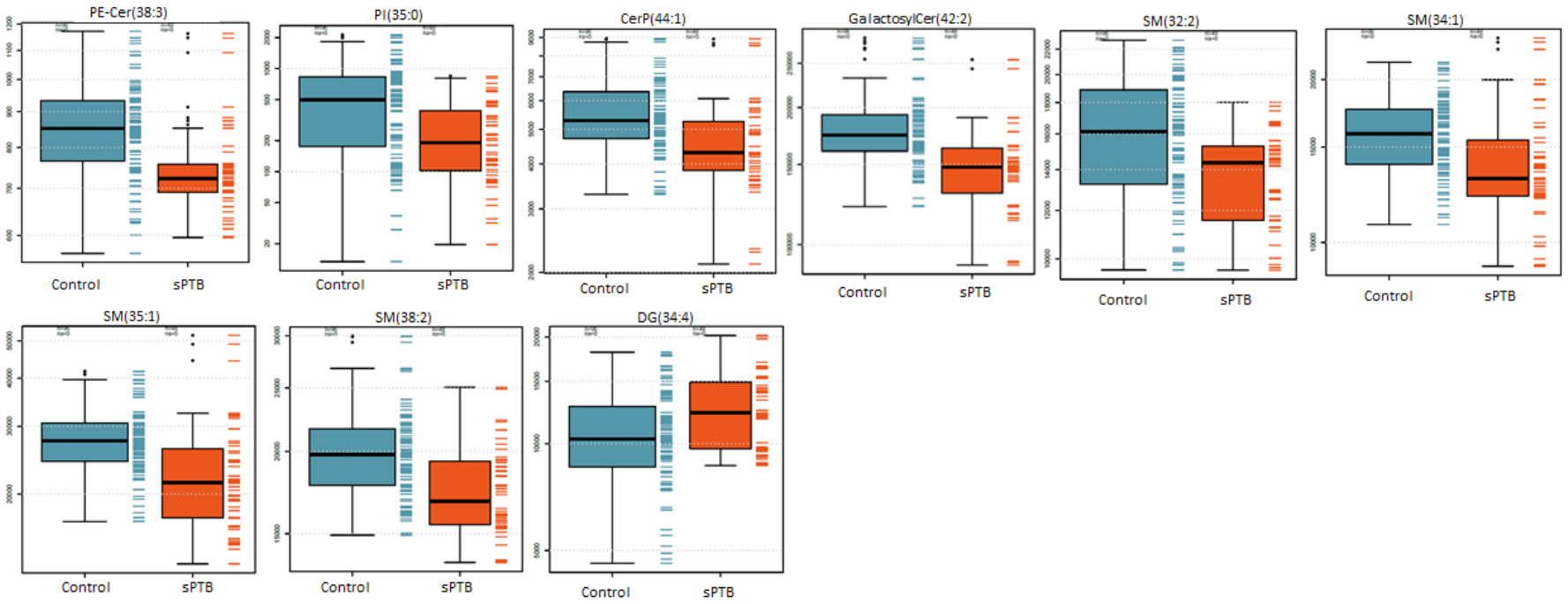



b
Control
Pilot3_Mar_Cont7b_Pos_2 883 (6.407)
SYNAPT G2-S#UEA009K
24-Mar-2017
2: TOF MS ES+
2.20e4
Relative intensity
m/z
[M+H]+
756.6105 21988
782.5351 18337
778.5980 15777
725.6185 10313
757.6196 9960
804.6110 8622
783.6369 7616
806.6307 7570
807.6353 2861
828.6118 2707
542.5372 6190
595.5161 6347
520.5475 5237
666.5364 5588
621.5317 3485
719.5167 3742
726.6184 3692
703.6333 3304
745.5442 2528
502.5392 1952
521.5539 2202
543.5341 1978
573.5328 1690
596.5179 1647
622.5333 1591
645.5388 1892
667.5398 1933
644.5964 1079
646.5374 718
701.5505 1014
704.6295 1245
727.6190 891
758.6292 1732
768.5933 982
794.6245 2029
816.6172 1478


c
Control
Pilot3_Mar_Cont7b_Pos_2 1244 (9.021)
SYNAPT G2-S#UEA009K
24-Mar-2017
1: TOF MS ES+
3.71e5
Relative intensity
799.7396
370968
773.7212
309440
800.7399
185807
[M+H-H2O]+
821.7148
65468
848.7250
37386
900.7643
16775
116.1498
25430
m/z
Pilot3_Mar_Cont7b_Pos_2 1244 (9.025)
2: TOF MS ES+
8.72e5
184.0868
872463
86.1016
63796
125.0068
39536
185.0891
36680
795.6955
27980
821.7148
28024
d

a: MS1 (top) and MS2 (bottom) of sphingomyelin SM(34:1) (HMDB0013464); **b:** MS2 of SM(34:1) zoom in range 490-830 m/z; **c:** MS1 (top) and MS2 (bottom) of phosphatidylcholine PC(42:4) (HMDB0008224); **d:** MS2 of PC(42:4) zoom in range 814-865 m/z


PC(35:1)
PC(40:5)
PC(42:4)
PC(42:5)
PC(O-38:4)
PC(O-42:4)
PC(O-44:6)
PC(P-36:1)
PC(P-38:3)
PC(P-40:3)
PC(P-40:4)
PC(P-40:5)
PE(38:0)
PE(38:1)
PE(38:2)
PE(40:3)
PE(40:4)
PE(42:6)
Control
sPTB



Box plots showing the normalised intensity of features significantly altered (adjusted p-value < 0.05) in spontaneous preterm birth (sPTB) group (orange box) compared to control group (blue box), from semi-targeted UPLC-MS analysis of SCOPE Cork (cases n=16, controls n=32). Segments represent observations from all triplicate injections.

Appendix D

Supplementary information for Chapter 5

Lipid concentrations of 7 calibrators prepared from SPLASH® Lipidomix®.

Calibrator (ng/mL)	DG (15:0/18:1-d_7)	TG (15:0/18:1-d_7/15:0)	PE (15:0/18:1-d_7)	PC (15:0/18:1-d_7)
7	1000	6000	602	17114
6	500	3000	301	8557
5	100	600	60	1711
4	20	120	12	342
3	2.0	12	1.2	34.2
2	0.5	3.0	0.3	8.6
1	0.07	0.4	0.04	1.1

Supplementary methods information

Recovery and repeatability of the deuterium labelled phosphatidylcholine (PC 15:0/18:1-d_7), phosphatidylethanolamine (PE 15:0/18:1-d_7), diacyglycerol (DG 15:0/18:1-d_7) and triacyglycerol (TG 15:0/18:1-d_7/15:0) were determined by spiking a 1:50 dilution of SPLASH® Lipidomix® into the 15 and 20 weeks' QC samples plasma samples (see Table bellow). The linear calibration curves were used for quality control purposes, with batch acceptance criteria of $\pm$30 % accuracy of 4 out of 6 QC samples within each batch.

Nominal concentrations were assigned to the native targets ((PC (15:0/22:6), PE (P-16:0/20:5), DG (14:0/18:1), Cer (m18:0/22:0), TG (18:3/18:2/18:3)) in the calibrator plasma matrix based on the calibration curve of the labelled analyte of each

respective lipid class, e.g. the concentration of PC (15:0/22:6) was assigned a nominal value using the calibration curve of PC 15:0/18:1-d_7, etc. (the DG was used to assign a nominal concentration for the ceramide as there was no labelled ceramide available). Semi-quantification of the samples was achieved using Average Response Factor in TargetLynx™ of the nominally assigned target concentration. True quantitation of the target lipids was not feasible due to the lack of availability of authentic standards.

Recovery and repeatability of deuterium labelled lipids from SPLASH® Lipidomix® spiked into quality control (QC) samples created from pooled plasma collected at 15 and 20 weeks of gestation (QC15 and QC20).

Lipid	**QC**	**Concentration (ng/mL)**	**Intra-day CV (%) n=3**	**Inter-day CV (%) n=8**	**Recovery (%)**
PC (15:0/18:1-d_7)	15	3012	1.9-24.9	8.6	4.6
	20	3012	2.5-12.2	9.1	-1.6
PE (15:0/18:1-d_7)	15	106	0.7-26.2	12.7	4.4
	20	106	2.8-13.9	8.9	1.5
DG (15:0/18:1-d_7)	15	176	1.8-20.3	8.1	16.5
	20	176	6.7-24.3	10.6	18.7
TG (15:0/18:1-d_7/15:0)	15	1056	0.1-24.6	21.8	-15.1
	20	1056	2.0-12.4	24.9	-23.5

PC = phosphatidylcholine; PE = phosphatidylethanolamine; DG = Diacylglycerol; TG = Triacylglycerol

Relative concentrations and measurement repeatability of five targeted lipids from plasma pooled from maternal samples collected at 15 and 20 weeks of gestation.

Lipid	Weeks of gestation	Relative concentration ±SD (ng/mL)	Intra-day CV (%) n=3 per day	Inter-day CV (%) n=8
PC (15:0/22:6)	15 20	3690 ±120 3638 ±94	0.7-2.7 0.6-3.7	3.2 2.6
PE (P-16:0/20:5)	15 20	34 ±5 34 ±1	0.6-9.7 1.7-12.8	5.4 3.6
DG (14:0/18:1)	15 20	309 ±29 401 ±22	3.2-18.4 3.1-14.4	9.2 5.4
Cer (m18:0/22:0)	15 20	209 ±98 290 ±120	1.3-15.1 2.2-23.2	46.9 39.2
TG (18:3/18:2/18:3)	15 20	3260 ±190 3740 ±130	1.9-7.2 1.8-6.0	5.8 3.4

PC = phosphatidylcholine; PE = phosphatidylethanolamine; DG = Diacylglycerol; TG = Triacylglycerol

Putatively annotated differential features in spontaneous preterm birth cases relative to full term birth controls from the discovery analysis.

Chemical class	Lipid Identification*	Normalised abundance Fold Change**	change relative to control	eBayes p value**
Glycerolipid	DG(38:3)	1.13	up	2.7E-02
Glycerolipid	DG(30:1)	1.83	up	9.9E-03
Glycerolipid	DG(32:0)	1.24	up	1.1E-02
Glycerolipid	**DG(32:1)**	**1.33**	**up**	**5.8E-03**
Glycerolipid	DG(32:2)	1.17	up	7.1E-03
Glycerolipid	DG(34:0)	1.36	down	7.6E-03
Glycerolipid	DG(34:1)	1.19	up	4.0E-03
Glycerolipid	DG(34:2)	1.15	up	8.6E-03
Glycerolipid	DG(34:3)	1.27	up	4.8E-04
Glycerolipid	DG(36:1)	1.18	up	2.8E-04
Glycerolipid	DG(36:2)	1.16	down	1.9E-03
Glycerolipid	DG(36:3)	1.13	up	1.1E-02
Glycerolipid	DG(36:4)	1.37	down	1.0E-03
Glycerolipid	DG(38:2)	1.18	down	1.3E-02
Glycerolipid	DG(38:4)	1.12	down	4.1E-03
Glycerolipid	DG(38:5)	1.16	up	1.3E-02
Glycerolipid	DG(40:4)	1.26	up	1.5E-02
Glycerolipid	DG(40:5)	1.31	down	2.3E-02
Glycerolipid	MG(16:0)	1.32	up	5.1E-03
Glycerolipid	MG(18:1)	1.14	up	7.3E-03
Glycerolipid	MG(20:4)	1.12	down	9.4E-03
Glycerolipid	MG(24:1)	1.21	down	1.5E-02
Glycerolipid	TG(48:2)	1.19	down	8.2E-03
Glycerolipid	TG(54:9)	1.20	down	2.2E-02
Glycerolipid	TG(58:9)	1.12	down	2.6E-02
Glycerolipid	TG(62:12)	1.12	down	1.7E-02
Glycerolipid	TG(62:13)	1.27	down	1.9E-02
Glycerolipid	TG(62:7)	1.14	down	3.9E-03
Glycerolipid	TG(68:9)	1.11	down	1.7E-02
Glycerolipid	**TG(54:8)**	**1.48**	**up**	**5.3E-03**
Glycerophospholipid	LPC(16:0)	1.11	down	5.7E-03
Glycerophospholipid	LPC(18:0)	1.23	down	2.5E-02
Glycerophospholipid	LPC(18:2)	1.13	down	2.1E-02
Glycerophospholipid	LPC(22:6)	1.15	down	9.4E-03
Glycerophospholipid	LPE(16:1)	1.25	up	2.0E-02
Glycerophospholipid	LPE(20:4)	1.15	down	1.7E-02
Glycerophospholipid	LPE(22:6)	1.21	down	1.7E-03
Glycerophospholipid	**PC(37:6)**	**1.63**	**down**	**1.2E-02**
Glycerophospholipid	PC(16:1/P-16:0)	1.25	down	1.2E-04
Glycerophospholipid	PC(20:3/P-18:0)	1.24	down	2.7E-03
Glycerophospholipid	PC(20:4/P-18:1)	1.23	down	4.3E-06
Glycerophospholipid	PC(22:5/P-18:0)	1.18	down	3.4E-05

Glycerophospholipid	PC(29:0)	1.27	up	1.2E-04
Glycerophospholipid	PC(32:1)	1.19	down	2.0E-04
Glycerophospholipid	PC(33:0)	1.11	up	1.3E-02
Glycerophospholipid	PC(34:2)	1.37	down	3.8E-03
Glycerophospholipid	PC(34:3)	1.25	down	2.4E-05
Glycerophospholipid	PC(34:4)	1.24	down	1.9E-04
Glycerophospholipid	PC(36:5)	1.17	down	4.4E-04
Glycerophospholipid	PC(37:6)	1.21	down	8.3E-05
Glycerophospholipid	PC(38:5)	1.17	down	2.0E-02
Glycerophospholipid	PC(38:7)	1.08	down	1.7E-02
Glycerophospholipid	PC(40:5)	1.14	down	5.4E-04
Glycerophospholipid	PC(40:6)	1.20	down	5.3E-04
Glycerophospholipid	PC(40:7)	1.14	down	1.1E-02
Glycerophospholipid	PC(42:7)	1.17	down	1.5E-04
Glycerophospholipid	PC(44:10)	1.09	down	2.8E-03
Glycerophospholipid	PC(44:6)	1.22	down	5.2E-04
Glycerophospholipid	PC(46:6)	1.17	down	3.2E-03
Glycerophospholipid	PC(o-16:0/14:1)	1.78	down	4.4E-04
Glycerophospholipid	PC(P-16:0/16:0)	1.12	down	3.6E-03
Glycerophospholipid	PC(P-16:0/18:3)	1.17	down	6.5E-03
Glycerophospholipid	PC(P-16:0/24:1)	1.09	down	1.0E-02
Glycerophospholipid	PC(P-18:0/14:1)	1.25	down	4.2E-05
Glycerophospholipid	PC(P-18:0/22:4)	1.16	down	7.8E-04
Glycerophospholipid	PC(P-18:0/22:6)	1.19	down	5.1E-03
Glycerophospholipid	PC(P-18:1/14:0)	1.22	down	2.5E-04
Glycerophospholipid	PC(P-18:1/14:1)	1.40	down	1.5E-05
Glycerophospholipid	PE(15:0/20:1)	1.25	down	7.6E-03
Glycerophospholipid	PE(18:0/P-16:0)	1.08	down	1.7E-02
Glycerophospholipid	PE(18:0/P-18:1)	1.13	down	1.5E-03
Glycerophospholipid	PE(22:5/P-18:0)	1.13	down	4.6E-03
Glycerophospholipid	PE(22:6/P-18:0)	1.15	down	1.5E-04
Glycerophospholipid	PE(22:6/P-18:1)	1.13	down	2.6E-03
Glycerophospholipid	PE(32:1)	1.18	up	9.4E-03
Glycerophospholipid	PE(32:2)	1.26	up	7.1E-04
Glycerophospholipid	PE(34:0)	1.27	up	1.1E-02
Glycerophospholipid	PE(34:1)	1.12	up	2.2E-03
Glycerophospholipid	PE(34:2)	1.21	up	1.0E-03
Glycerophospholipid	PE(34:2)	1.10	up	1.4E-02
Glycerophospholipid	PE(34:3)	1.24	up	5.1E-04
Glycerophospholipid	PE(36:1)	1.14	up	8.2E-04
Glycerophospholipid	PE(36:2)	1.08	up	6.9E-03
Glycerophospholipid	PE(36:3)	1.09	up	3.9E-03
Glycerophospholipid	PE(38:4)	1.10	down	2.4E-03
Glycerophospholipid	PE(38:4)	1.21	up	1.5E-02
Glycerophospholipid	PE(40:6)	1.11	down	3.5E-04
Glycerophospholipid	PE(42:0)	1.13	down	6.3E-05
Glycerophospholipid	PE(44:1))	1.16	down	1.7E-03
Glycerophospholipid	PE(44:2)	1.13	down	2.4E-03
Glycerophospholipid	PE(52:4)	1.17	down	9.8E-03
Glycerophospholipid	PE(P-16:0/18:2)	1.19	down	4.2E-03

Glycerophospholipid	PE(P-16:0/20:4)	1.14	down	2.5E-04
Glycerophospholipid	**PE(P-16:0/20:5)**	**1.33**	**down**	**7.4E-04**
Glycerophospholipid	PE(P-18:1/15:0)	1.14	down	2.3E-02
Glycerophospholipid	PE(P-18:1/16:0)	1.23	down	6.8E-04
Glycerophospholipid	PE(P-18:1/18:1)	1.20	down	3.0E-04
Glycerophospholipid	PG(15:0/22:2)	1.18	down	1.1E-02
Glycerophospholipid	PS(34:0)	1.18	down	2.5E-02
Glycerophospholipid	PS(38:1)	1.09	down	5.7E-03
Glycerophospholipid	PS(38:2)	1.17	down	1.1E-03
Glycerophospholipid	PS(39:2)	1.13	down	5.5E-04
Glycerophospholipid	PS(39:3)	1.11	down	2.2E-02
Glycerophospholipid	PS(40:1)	1.15	down	2.6E-05
Glycerophospholipid	PS(41:2)	1.17	down	1.5E-04
Glycerophospholipid	PS(43:2)	1.20	down	1.5E-02
Glycerophospholipid	PS(O-20:0/17:2)	1.18	down	7.0E-03
Sphingolipid	Cer(d18:0/22:0)	1.19	up	8.9E-04
Sphingolipid	Cer(d18:1/22:0)	1.14	up	5.7E-04
Sphingolipid	Cer(d18:1/22:1)	1.21	up	2.9E-03
Sphingolipid	**Cer(m18:0/22:0)**	**1.31**	**up**	**8.5E-03**
Sphingolipid	Cer(m18:0/24:0)	1.19	up	2.7E-02
Sphingolipid	Cer(m18:1/24:0)	1.16	up	6.8E-03
Sphingolipid	Cer(t18:0/22:0(2OH))	1.10	up	2.4E-02
Sphingolipid	Galabiosylceramide (d18:1/18:1)	1.07	down	2.2E-02
Sphingolipid	Galabiosylceramide (d18:1/20:0)	1.13	down	7.0E-03
Sphingolipid	Galabiosylceramide(d18 :1/16:0)	1.19	down	2.6E-04
Sphingolipid	GlcCer(d18:2/20:1)	1.12	down	1.4E-03
Sphingolipid	SM(d18:2/18:1)	1.13	down	1.4E-02
Sterol lipid	5b-Cholestane-3a,7a,12a,23R,25-pentol	1.31	up	2.0E-02
Sterol lipid	6-Deoxocastasterone	1.18	up	1.3E-02
Sterol lipid	CE(22:6)	1.19	down	4.1E-04

*Lipids putatively annotated using Progenesis software resulting from Human Metabolome Database and LIPIDMAPS searches, ** eBayes adjusted for FDR at 5 % p-value <0.05, statistic provided by SQU4RE. Bold: Features selected for targeted analysis. DG = Diacylglycerol; TG = Triacylglycerol; PC = phosphatidylcholine; PE = phosphatidylethanolamine; Cer = Ceramide.