

Association of newborn screening analytes with type of delivery among preterm and term births

Jessica Yau

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial fulfillment of the requirements for the degree of Master of Science in Epidemiology

School of Epidemiology and Public Health

Faculty of Medicine

University of Ottawa

© Jessica Yau, Ottawa, Canada, 2018

ABSTRACT

Introduction: Several factors have been observed to influence the value of newborn screening analytes (NBS) and should be adjusted for in the interpretation of blood spot samples. The thesis aimed to examine the association of NBS with 1) mode of delivery among term infants and 2) clinical subtypes (i.e., spontaneous onset of labour) of birth among preterm infants.

Methods: A retrospective population-based cross-sectional study design was employed. A multivariable logistic regression model was used to examine associations between NBS and mode of delivery among term infants and subtypes among preterm infants.

Results: 1) Metabolic profiles of infants born by planned cesarean delivery differ from those born by vaginal delivery following spontaneous onset of labour and 2) Metabolic profiles of preterm infants did not differ by clinical subtype.

Conclusions: Our findings conclude that mode of delivery is an important covariate to consider in future modelling studies, but the inclusion of preterm birth subtypes is less compelling.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Steven Hawken. I am so lucky to have been under the guidance of such a brilliant mind, filled with biostatistics knowledge. Thank you for always challenging me to work through SAS syntax by myself but also lending a helping hand whenever I needed it!

I am incredibly grateful for Dr. Deshayne Fell, my co-supervisor, I always feel like I am ready to tackle the world after your encouraging words and constant support. Thank you for providing such a conducive environment for my growth professionally, as well as, personally.

A huge thank you to my committee members: Dr. Beth Potter, Dr. Pranesh Chakraborty, Dr. Kumanan Wilson and a collaborator to my thesis: Dr. Mark Walker. Their insight, guidance and technical expertise were invaluable in completing this degree.

To Candis Lepage, my number one rock throughout this degree. Thank you for doing life with me these past two years and for introducing Lisa into my life. I couldn't have done this without you!

Finally, I would like to thank my family and friends for providing such a solid foundation of support and love. You have helped me in ways you'll never know.

STATEMENT OF FINANCIAL SUPPORT

It should be acknowledged that this thesis was made possible by funding from:

- 1) A Grand Challenges Explorations grant from the Bill & Melinda Gates Foundation [OPP1141535].
- 2) A part-time studentship at the Children's Hospital Eastern Ontario within Dr. Fell's research team.
- 3) Ontario Graduate Scholarship, a non-renewable scholarship awarded \$15,000 for 12 months (May 1st, 2017 to May 1st, 2018).
- 4) University of Ottawa Excellence Scholarship.

Table of Contents

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
STATEMENT OF FINANCIAL SUPPORT.....	iv
Table of Contents.....	v
List of Tables.....	viii
List of Figures.....	ix
List of Acronyms and Abbreviations.....	xi
CHAPTER 1. INTRODUCTION.....	1
1.1 Background.....	1
1.2 Research Objectives.....	3
CHAPTER 2. LITERATURE REVIEW.....	4
2.1 High Dimensional Biology.....	4
2.2 High Dimensional Biology in Maternal-Fetal Medicine.....	4
2.3 Newborn Screening.....	5
2.3.1 Newborn Screening Ontario.....	6
2.3.2 Newborn Screening Analytes.....	7
2.3.3 Interpretation of Newborn Screening Analytes.....	8
2.4 Mode of Delivery.....	9
2.4.1 Cesarean Delivery.....	9
2.4.2 Vaginal Delivery.....	11
2.5 Preterm Birth.....	12
2.5.1 Spontaneous Preterm Birth.....	13
2.5.2 Medically-indicated Preterm Birth.....	14
2.6 Summary.....	14
CHAPTER 3. METHODOLOGY.....	16
3.1 Study Design.....	16
3.2 Data Sources.....	16
3.2.1 Data Linkage, Data Access, and Privacy Considerations.....	18
3.3 Study Population.....	18
3.4 Data Cleaning and Preparation.....	22
3.5 Dependent Variables.....	24
3.6 Independent Variables.....	25

3.7 Statistical Analyses	26
3.8 Sensitivity Analyses	32
3.9 Ethical Considerations.....	33
CHAPTER 4. RESULTS.....	34
4.1 Objective 1	34
4.1.1 Characteristics of the Study Population.....	34
4.1.2 Heat Map Analysis.....	37
4.1.3 Partial Correlation Analysis	39
4.1.4 Overall Model Performance.....	39
4.1.5 Overall Model Performance from Sensitivity Analyses	46
4.2 Objective 2	49
4.2.1 Characteristics of the Study Population.....	49
4.2.2 Heat Map Analysis.....	53
4.2.3 Partial Correlation analysis	55
4.2.4 Overall Model Performance	55
4.2.5 Overall Model Performance from Sensitivity Analyses	62
CHAPTER 5. DISCUSSION.....	66
5.1 Statement of Principal Findings	66
5.2 Interpretation of Findings	67
5.3 Strengths	70
5.4 Limitations.....	72
5.5 Conclusions	75
REFERENCES.....	77
APPENDIX A: Certificates for Ethical Approval.....	86
APPENDIX B: Letters of Support.....	91
APPENDIX C: Histograms for Objective 1	94
APPENDIX D: Histograms for Objective 2.....	100
APPENDIX E: Boxplots for Objective 1	106
APPENDIX F: Boxplots for Objective 2	112
APPENDIX G: Partial Correlation Analysis Results for Objective 1.....	118
APPENDIX H: Forest Plots for Sensitivity Analyses for Objective 1.....	120
APPENDIX I: Partial Correlation Analysis for Objective 2.....	128

APPENDIX J: Forest Plots for Sensitivity Analyses for Objective 2 Supplemental Figure J.1a Individual analytes and clinical variables in the final model of preterm, including those with transfusions 130

List of Tables

Table 3.1 Newborn screening analytes available in Newborn Screening Ontario database	28
Table 3.2. List of intrapartum and obstetrical complications and maternal health problems leading to exclusion of infants born at term that were not considered to be delivered from low-risk women	31
Table 4.1 Baseline characteristics of full, training and validation datasets for singleton term live births in Ontario between 2010 and 2012 by mode of delivery	48
Table 4.2 C-statistics for discrimination performance for Objective 1	52
Table 4.3 C-statistic for sensitivity analyses for Objective 1	58
Table 4.4 Baseline characteristics of singleton preterm live births in Ontario between 2008 and 2012 by preterm birth subtype	64
Table 4.5 C-statistics for discrimination performance for Objective 2	68
Table 4.6 C-statistic for sensitivity analyses for Objective 2	74

List of Figures

Figure 3.1 Graphical presentation of dependent variable for Objective 1 and 2	36
Figure 4.1 Study flow diagram for Objective 1	46
Figure 4.2 Correlation heatmap of analytes for infants with planned cesarean delivery	50
Figure 4.3 Correlation heatmap of analytes for infants with vaginal delivery following spontaneous onset of labour	50
Figure 4.4 Decile calibration plot for term infants	53
Figure 4.5a Individual analytes and clinical variables in the final model for term infants	56
Figure 4.5b Analyte ratios in the final model for term infants	57
Figure 4.6 Decile calibration plots for early-term infants	59
Figure 4.7 Decile calibration plots for full-term infants	59
Figure 4.8 Decile calibration plots for infants who had blood spot collection within recommended window	60
Figure 4.9 Decile calibration plots after inclusion of remaining ranked analyte ratios until the maximum number of parameters was reached	60
Figure 4.10 Study flow diagram for Objective 2	63
Figure 4.11 Correlation heatmaps for preterm infants born following PPROM	66
Figure 4.12 Correlation heatmaps for preterm infants born following spontaneous onset of labour	66
Figure 4.13 Decile calibration plots for preterm infants	69
Figure 4.14a Individual analytes and clinical variables in the final model for preterm infants	72
Figure 4.14b Analyte ratios in the final model for preterm infants	73
Figure 4.15 Decile calibration plots for preterm infants, including those who underwent transfusions	75
Figure 4.16 Decile calibration plots for preterm infants born between 2010 and 2012	75
Figure 4.17 Decile calibration plots for preterm infants, with the inclusion of TPN in the model	76
Figure 4.18 Decile calibration plots after inclusion of remaining ranked analyte ratios until the	76

maximum number of parameters was reached

List of Acronyms and Abbreviations

AUC – Area under the curve

BORN – Better Outcomes Registry & Network

CD – Cesarean delivery

CHEO – Children’s Hospital of Eastern Ontario

HDB – High dimensional biology

NSO – Newborn Screening Ontario

NBS – Newborn screening

PPROM – Preterm premature rupture of membranes

ROC – Receiver operating curve

TPN – Total parenteral nutrition

CHAPTER 1. INTRODUCTION

1.1 Background

Over recent decades, high dimensional biology (HDB) has emerged as an approach to understanding many complex biological processes and disorders.¹ HDB refers to the study of changes in the genome, transcriptome, proteome, or metabolome to understand the pathophysiology of disease.^{2,3} In obstetrics, using samples of biological fluids obtained from the mother and newborn such as amniotic fluid, urine, saliva, and blood, biochemical markers have been able to provide a rich source of information to better understand the complex processes related to pregnancy such as parturition and the biological mechanisms associated with several disorders of pregnancy, including preterm birth.^{2,4,5}

In Ontario, Canada, virtually every child undergoes newborn screening.^{6,7} The Newborn Screening Ontario (NSO) program collects heel prick blood spot samples from newborns shortly after birth and screens each infant for 29 rare, but treatable, conditions with the use of a panel of screening analytes. The purpose of this population-wide screening program is to identify newborns with rare diseases at an early stage in order to facilitate early intervention and reduce morbidity and mortality.⁸ Collectively, these diseases will be identified in approximately 200 out of 145,000 children born in Ontario each year.⁷ Factors known to influence the values of the newborn screening analytes, potentially leading to false negative and false positive rates,^{9,10} including gestational age and birth weight.¹⁰⁻¹² These factors are taken into consideration when collecting and interpreting the results of newborn screening analysis. Several environmental and physiological factors such as season of birth, gestational diabetes and mode of delivery are known to affect long-term maternal and neonatal health outcomes.^{10,13,14} However, their influence on the concentrations of analytes obtained through newborn screening is unclear.

Preterm birth, defined as birth prior to 37 completed weeks of gestational age, is a leading cause of infant and childhood mortality and morbidity.^{15,16} Worldwide, an estimated 15 million newborns are born preterm and this number is increasing.¹⁵ Despite extensive research,¹⁷ the etiology of preterm birth remains unclear but is known to be multifactorial.⁵ Approximately 30% of preterm births are medically-indicated and the remaining 70% occur spontaneously. While spontaneous and medically-indicated preterm births may share some clinical features,¹⁸ the latter are mainly performed for maternal or fetal indication,¹⁹ whereas spontaneous preterm births occur more frequently due to pathways involving infection and inflammation.^{2,20}

This thesis project used data from Ontario's province-wide newborn screening program which has been linked with the provincial birth registry, providing an opportunity to explore whether newborn screening analyte profiles (e.g., levels of amino acids, enzymes, and endocrine markers in the blood spot analyzed as part of the screening program) are associated with type of delivery among preterm and term births. When used beyond their traditional application, such as in postnatal gestational age estimation²¹ and predictive models for other neonatal morbidities,²² several studies have demonstrated associations between newborn screening analytes and maternal and neonatal outcomes. However, a thorough investigation into the influences of environmental and biological factors, including mode of delivery and subtypes of preterm birth, has not yet been undertaken to understand the variability in newborn screening analytes. The purpose of the thesis project was to understand the associations between newborn screening analytes with type of delivery among preterm and term births.

Specifically, it aims to provide a post hoc description of mode of delivery among term births and subtypes of birth among preterm births in relation to newborn screening analytes to determine if when used beyond their traditional application (i.e., predictive modelling studies) whether they should be taken into consideration as potential confounders. This thesis project is not for causal or predictive

purpose and is not aiming to use newborn screening analytes to predict type of delivery outcomes among preterm and term births. However, it can inform etiologic or causal understanding by generating hypotheses on how type of delivery is associated with newborn screening analytes.

1.2 Research Objectives

The overall aim of this master's research was to investigate the association between type of delivery (as defined below) among preterm and term births with newborn screening analyte values, using newborn screening data linked with provincial birth registry data. The two specific objectives of the thesis project were to determine whether there is an association between newborn screening analyte profiles and:

- i. Mode of delivery (planned cesarean delivery or vaginal delivery following spontaneous onset of labour) among infants born at term gestation to low-risk women
- ii. Birth subtypes (delivery following spontaneous onset of labour or preterm premature rupture of membranes) among infants born at preterm gestation

CHAPTER 2. LITERATURE REVIEW

2.1 High Dimensional Biology

High-dimensional biology (HDB) refers to the study of the genome, transcriptome, proteome, or metabolome in biological samples with the intent of understanding the pathophysiology of diseases.^{3,23} HDB encompasses the “omics” sciences, which refers to the study of entities in aggregate. With the development of HDB over the past two decades, the simultaneous screening of thousands of genes, gene products, proteins and metabolites from small samples of tissue or body fluid has become possible.²⁴

2.2 High Dimensional Biology in Maternal-Fetal Medicine

Increasingly, the application of HDB has been used in maternal-fetal medicine to provide insight on the biological processes involved in the process of parturition and the complex disorders of pregnancy, such as preterm birth.^{3,24–29} Maternal biological fluids including cervicovaginal fluid, amniotic fluid, serum, blood, urine, and saliva have been studied as they provide rich sources of proteins and metabolites that vary in concentration in response to pregnancy itself, as well as in response to adverse pregnancy states.^{4,30–32} The majority of the literature regarding HDB involves maternal bodily fluids.^{4,33–37} Several systematic reviews have been conducted to summarize studies of biomarkers that predict spontaneous preterm birth. However, few biomarkers have shown clinical usefulness. Three systematic reviews have identified cervicovaginal fetal fibronectin – a glycoprotein that is a marker of choriodecidual disruption⁵ – as the biomarker with the highest strength of association with spontaneous preterm birth.^{32,33,37} However, in another systematic review, the authors concluded that due to methodological differences in study designs, timing of blood collection or other samples and study populations, a meta-analysis was not possible and

thus could not conclude if a single biomarker will be able to predict spontaneous preterm birth or to understand the underlying pathophysiological mechanisms associated with this condition.³⁶

There is a paucity of research on spontaneous preterm birth that has targeted biological samples obtained from the newborn. One prospective study examined both maternal serum and newborn cord blood to identify markers of infection (e.g., interleukin 1 β and tumor necrosis factor α) to assess whether they could predict spontaneous preterm birth;³⁸ however, the cord blood sample was used to confirm fetal exposure to pro-inflammatory cytokines rather than as an independent sample for primary analysis.³⁸ This study echoed the conclusions made by the systematic review described previously in concluding that identification of a single predictive marker associated with preterm birth is unlikely.³⁸ Rather, identifying multiple markers will be necessary, as the etiology of preterm birth is known to be multifactorial.³⁸ While the objective of most previous research has been the prediction of preterm birth, one objective of the current thesis project is to understand the associations between subtypes of preterm birth with newborn screening analytes.

2.3 Newborn Screening

Newborn screening is a population-wide screening program that is conducted in approximately 64 countries worldwide.⁸ Dried blood spot samples are collected shortly after birth to identify rare, treatable disorders that usually show no symptoms in the newborn period but where early identification and initiation of treatment is important for achieving reductions in morbidity and mortality.³⁹ There is considerable variability in the disorders included in screening programs from country to country, and even across jurisdictions within the same country.⁸ This can be due to a variety of factors including prevalence of specific disorders in a geographic region, availability of

treatment for the targeted disorder, lack of certainty in the available evidence regarding the effectiveness of early intervention, and clinical practice guidelines.⁸

Within Canada, there is no national newborn screening program. Each province and territory screens for a different set of diseases.⁴⁰ The number of conditions screened ranges substantially from 6 in Manitoba to over 30 in Saskatchewan.⁴¹ Phenylketonuria (PKU), congenital hypothyroidism (CH), and medium chain acyl-CoA dehydrogenase (MCAD) deficiency are the only three conditions that are screened in all programs in Canada.⁴¹

2.3.1 Newborn Screening Ontario

In Ontario, Newborn Screening Ontario (NSO) coordinates newborn screening which is based through a central laboratory at the Children's Hospital of Eastern Ontario and is funded by the provincial Ministry of Health and Long-Term Care. Heel prick blood samples are collected from all newborns within 24-72 hours of birth, from which levels of various analytes are measured.⁶ These analytes include acyl-carnitines, amino acids, endocrine markers, enzyme markers, and measurement of fetal (HbF) and adult (HbA) hemoglobin levels.^{21,40,42-45} Using the quantitative values of the analytes, NSO screens for treatable diseases including metabolic diseases (e.g., maple syrup urine disease), endocrine diseases (e.g., congenital hypothyroidism), sickle cell disease, cystic fibrosis, and severe combined immune deficiency.⁴⁶

Most of the screening analytes are measured using tandem mass spectrometry (MS/MS). A positive finding from MS/MS could point to several different disorders and is not diagnostic. Infants who

screen positive are referred to specialists at the Newborn Screening Regional Treatment Centres in the province where additional analytes are collected for confirmatory diagnostic testing.

Identification of newborns with a disorder at a young age can help prevent health problems, as treatment and management for the disease is provided. Currently, the province screens for 29 rare disorders. Collectively, these diseases will affect approximately 200 out of 145,000 infants born in Ontario each year. Over 1,600 infants have been diagnosed with a screened disease since the province-wide screening program was established in 2006.^{6,7}

2.3.2 Newborn Screening Analytes

The newborn screening analytes collected by Newborn Screening Ontario include acyl-carnitines, amino acids, HbF and HbA levels, endocrine markers and enzyme markers. Acyl-carnitines and amino acids are markers for metabolic disorders of fatty acid oxidation, organic acid disorders and amino acidopathies including, for example, medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, phenylketonuria and maple syrup urine disease. These are inherited disorders that affect the infant metabolism and are measured using MS/MS.⁴⁷

Hemoglobinopathies are inherited blood disorders that can cause abnormal structure or volume of hemoglobin (Hb). Measuring variants of Hb levels allows these disorders to be detected. Sickle cell disease is the primary hemoglobin disorder screened by the presence of Hb S. However, beta-thalassemia major, which is characterized by abnormal volume of hemoglobin, is also a target of newborn screening.⁴⁷

Endocrine markers include thyroid stimulating hormone (TSH) and 17-hydroxyprogesterone (17-OHP). NSO screens for two endocrine disorders: congenital hypothyroidism (CH) and congenital adrenal hyperplasia (CAH). These disorders are characterized by elevated levels of TSH and 17-OHP in CH and CAH, respectively.^{48,49}

The enzyme markers that are measured by NSO include: biotinidase (BIOT), galactose-1-phosphate uridylyltransferase (GALT) and immunoreactive trypsinogen (IRT). Screening for galactosemia and biotinidase deficiency is completed by measuring BIOT and GALT by fluorometry. Using immunoassay, NSO screens for cystic fibrosis by measuring IRT, which is a pancreatic enzyme precursor and its concentration is elevated in babies with cystic fibrosis.⁴⁷

2.3.3 Interpretation of Newborn Screening Analytes

There are several factors that are taken into account for the use of dried blood spots for newborn screening analysis. Blood volume and hematocrit can affect the precision and accuracy of analyte measurements.⁵⁰ In addition, clinical and environmental factors including gestational age of the infant at birth, sex, birth weight and feeding method also influence newborn screening analyte levels as well as false-positive and false-negative screening rates.^{10,12,51-53} For example, higher levels of 17-hydroxyprogesterone have been observed in males compared to females and this correlated with more false positive results in one study.¹⁰ While these factors are adjusted for in the interpretation of newborn screening results, there are a number of environmental and physiological factors stemming from conditions, complications and events during pregnancy and delivery that could also impact newborn screening analyte profiles, but there has been a paucity of research on this to-date.

2.4 Mode of Delivery

One such factor that may be associated with newborn screening analyte profiles is the mode of delivery, which is well known to be associated with other pediatric health outcomes. For instance, children delivered by cesarean are more likely to have atopic disease,¹³ cumulative infectious diseases and recurrent cough compared to children born following vaginal delivery.⁵⁴ The exposure to maternal hormonal responses, intermittent contractile forces and oxidative stress (due to spontaneous rhythmic contractions)^{55,56} in vaginal delivery support multiple physiological changes including an increase in fetal cortisol and catecholamine.⁵⁷ In contrast, infants born by cesarean delivery have been reported to experience blunting of the postnatal rise in cortisol and have lower levels of cord catecholamine levels.^{57,58} The physiological transition from a fetus to a newborn differs depending on mode of delivery – findings which warrant further research of the associations between profiles of newborn screening analytes and mode of delivery.

2.4.1 Cesarean Delivery

Cesarean delivery rates have been increasing in Canada, from 6% to upwards of 25% between 1970 and 2017.^{59,60} There are many complex factors contributing to this overall increase, including changes in clinical practice, changes in maternal demographics, and changes in attitudes and beliefs about birth and perception of risk.^{59,61} A number of medical indications can warrant a cesarean delivery including a previous cesarean section, dystocia, breech presentation, fetal distress, fetal overgrowth, and pre-eclampsia.^{59,61} Specific to Objective 1, this thesis project focused on planned cesarean delivery and we have included three clinical scenarios that comprise

this definition including: repeat cesarean delivery, breech presentation, and maternal request which are discussed in further detail in the following sections.

Repeat Cesarean Section

Around the 1990s, vaginal birth after cesarean (VBAC) was proposed as an attempt to reduce cesarean section rates through a reduction in subsequent cesarean deliveries to women who already had a previous cesarean.⁶² However, as VBAC became more common, higher rates of morbidity and mortality in mothers and infants were observed including hypoxic ischemic encephalopathy, endometritis, and uterine rupture in comparison to elective repeat cesarean delivery.⁶³⁻⁶⁵ This led to a reluctance by physicians to recommend a trial of labour in women whose first delivery was by cesarean section, as rupture of the uterine scar during labour is a primary concern when VBAC is considered.⁶² While the relative risk of uterine rupture with a trial of labour is higher than in planned repeat cesarean deliveries, the absolute risk remains low.⁴⁷⁻⁵⁰ The choice of VBAC or planned repeat cesarean delivery after a prior cesarean delivery is influenced by a combination of factors including interactions between patients and physicians, past experiences with labour and birth, and beliefs and attitudes about childbirth.⁵⁹

Breech Presentation

The International Term Breech Trial, which enrolled 2088 women with a singleton fetus in breech presentation at term gestation from 121 centres in 26 countries, concluded that the risk of perinatal or neonatal mortality or morbidity was significantly lower in the planned cesarean arm

compared to the planned vaginal delivery arm, with no significant difference in the risk of maternal mortality or morbidity.⁶⁶ This study resulted in a major change in clinical practice and shifted the routine management of term breech infants to planned cesarean delivery.⁶⁷ The option of an external cephalic version to avoid breech presentation is possible; however, the efficacy of the procedure is variable.⁶⁸ While some health care providers may opt to vaginally deliver women with a breech presentation who meet specified clinical criteria, the management of a term breech delivery remains somewhat controversial.⁶⁸

Maternal Request

The rising rates of cesarean delivery can be also partly attributed to maternal request.^{69,70}

Cesarean delivery on maternal request is a planned cesarean delivery prior to the onset of labour and with no medical or obstetrical indications for the cesarean.⁷¹ Several reasons prompting the request of cesarean deliveries from mothers include maternal morbidities that are associated with vaginal delivery including urinary incontinence, anal sphincter damage as well as fear of pain from labour and vaginal birth.^{59,61} The desire to schedule birth has also prompted some women to request cesarean section.⁶³ Results from a survey by the World Health Organization found that cesarean sections without medical indications were associated with an increased risk of adverse maternal outcomes.⁷²

2.4.2 Vaginal Delivery

Although rates of cesarean deliveries have been increasing over recent decades, the majority of women still delivery vaginally. A vaginal delivery is the process in which the infant is delivered

through the vaginal canal. There are different types of vaginal deliveries including: a vaginal delivery following spontaneous onset of labour, a vaginal delivery following induction of labour, and an assisted vaginal delivery (aided by the use of forceps or vacuum). Specific to Objective 1, this project specifically focused on vaginal delivery following spontaneous onset of labour as most studies comparing mode of delivery used this type of vaginal delivery in comparison to planned cesarean delivery. A vaginal delivery following spontaneous onset of labour is a natural process in which the mother goes into labour without the use of drugs or procedures to induce labour.⁷³

2.5 Preterm Birth

Preterm birth is a leading cause of perinatal morbidity and mortality worldwide⁷⁴ and the most common cause of death in children under 5 years old.¹⁶ Globally, it was estimated that in 2010 approximately 14.9 million babies were born preterm, comprising approximately 11% of all live births.^{16,74} The rates of preterm birth vary from country to country, with the highest rates of upwards of 12% seen in Southeastern Asia, South Asia and Sub-Saharan Africa.¹⁶ In Canada, approximately 8.3% of live births were born preterm in 2014.⁶⁰ In addition to its contribution to mortality, preterm birth is associated with a risk of long-term sequelae for the newborn including lifelong effects from neurodevelopmental impairments such as cerebral palsy, impaired learning, as well as higher risk for chronic disease in adulthood.^{5,16}

Preterm birth is defined as delivery prior to 37 weeks of completed gestation. Unlike other adverse perinatal outcomes, it is defined only by the timing of birth, regardless of the clinical events that precede it.^{17,18,74} Despite being extensively studied, little progress has been made toward understanding the etiology of preterm birth or in developing preventive interventions. It is

becoming increasingly evident that our understanding of preterm birth has been complicated by the difficulty in defining and classifying preterm birth.^{17,75} Existing phenotypic classifications include subdivisions by gestational age, clinical presentation, and presumed pathophysiological pathways.¹⁷ In addition to the various methods of classification, preterm birth is known to have a multitude of heterogeneous risk factors that can involve medical, genetic, social, and environmental components.⁵ Based on clinical presentation, preterm birth can be classified into: spontaneous preterm birth and medically-indicated preterm birth. The two subtypes may share common etiologies,^{18,19} but are also thought to have distinct pathways leading to the same outcome.¹⁸

2.5.1 Spontaneous Preterm Birth

The biological pathways that trigger spontaneous preterm birth are largely unknown but are often associated with infection and inflammation.⁷⁶ Spontaneous preterm birth can be further broken down into those following the spontaneous onset of labour with intact membranes and those initiated by a preterm premature rupture of the membranes (PPROM).^{5,19} Approximately 40-45% of preterm births occur following spontaneous labour and 25-30% can be classified as PPRM.^{5,18,74} PPRM is defined as spontaneous rupture of the membranes at less than 37 weeks of completed gestation and at least one hour prior to the onset of labour.⁵ The cause of membrane rupture is unknown but it has been associated with decreased collagen content. Intrauterine infection has been known to stimulate the production of matrix metalloproteinases (MMPs), enzymes that work to break down collagen which can lead to membrane rupture.^{2,77}

2.5.2 Medically-indicated Preterm Birth

Up to 30% of preterm births are medically-indicated, which is defined as a preterm delivery initiated through induction of labour or cesarean delivery prior to any naturally occurring onset of labour or membrane rupture.⁷⁸ Medically-initiated delivery prior to 39 completed weeks of gestation has long been discouraged, as the risks to the neonate such as respiratory distress syndrome and intraventricular hemorrhage have been established.^{78,79} However, a number of maternal and fetal indications may warrant early timing of delivery, such as pre-eclampsia, intrauterine growth restriction, and placental abruption.⁷⁸ These indications that motivate medical intervention seem to share mechanisms including inflammation and vascular compromise, which are pathways that are also known to lead to spontaneous preterm birth.¹⁸ The decision on the timing of delivery should balance maternal and newborn risks as well as take into account practice environment, level of neonatal care availability, and patient preferences.⁸⁰

2.6 Summary

With the recent development of HDB and its increasing application in maternal-fetal medicine, the simultaneous screening of thousands of genes, gene products, proteins and metabolites from small samples of tissue or body fluid has been used to provide insight on the biological processes involved in the process of parturition and the complex disorders of pregnancy, such as preterm birth.^{3,24-29}

In Ontario, the NSO program collects blood samples from the newborn shortly after birth and screens each infant for 29 rare, but treatable, conditions with the use of a panel of screening analytes.

Several factors are known to influence the values of newborn screening analytes which could lead to false negative and false positive rates^{9,10} such as gestational age and birth weight.¹⁰⁻¹² This thesis project aimed to examine the association of newborn screening analytes and mode of delivery. The physiological transition from a fetus to a newborn differs depending on mode of delivery – findings which warrant further research of the associations between profiles of newborn screening analytes and mode of delivery. Understanding the association of delivery parameters, specifically mode of delivery, with the quantitative values of newborn screening analytes may improve our interpretation of newborn screening analyses.

Despite being extensively studied, the etiology of preterm birth remains unclear but is known to be multifactorial.⁵ Our understanding of preterm birth has been complicated by the difficulty in defining and classifying preterm birth.^{17,75} Existing phenotypic classifications include subdivisions by gestational age, clinical presentation, and presumed pathophysiological pathways.¹⁷ Based on clinical presentation, this thesis project aimed to determine whether there is any association between newborn screening analyte profiles and subtypes of preterm birth. Understanding the associations of single or multiple screening analytes with subtypes of preterm birth could help generate hypotheses about biological mechanisms underlying preterm birth or subtypes of preterm birth.

CHAPTER 3. METHODOLOGY

The initial sections of this Chapter (Section 3.1 to Section 3.9) describe the study methods that apply to both Objectives 1 and 2. Objective-specific methods are then described in later sections.

3.1 Study Design

A retrospective population-based cross-sectional study design was employed for both objectives of this thesis project. These two studies were an extension of an ongoing retrospective, population-based cohort study in the province of Ontario. Pursuing both objectives involved a secondary analysis of data collected at NSO and linked with health administrative data and the Better Outcomes & Registry Network (BORN)-Niday Perinatal database, which are housed at ICES.

3.2 Data Sources

The NSO database collects screening records for all infants in the province who have been tested for 29 rare conditions using a panel of screening analytes, obtained shortly after birth via heel prick blood spot. Of all the infants in Ontario, approximately 99.7% are screened by NSO.⁸¹ Using standardized biochemical assays, newborn screening analytes available at NSO include acyl-carnitines, amino acids, fetal-to-adult Hb level, endocrine and enzyme markers. A full list of available analytes can be found in Table 3.1. The NSO database also contains information on timing of the blood sample collection, method of feeding (e.g., breast, formula, or total parenteral nutrition) and transfusion of blood or blood products.

Table 3.1 Newborn screening analytes available in Newborn Screening Ontario database

Categories of Newborn Screening Analytes	Newborn Screening Analyte	Description of Analyte
Acyl-carnitines (n=31)	C0; C2; C3; C4; C5; C5:1; C6; C8; C8:1; C10; C10:1; C12; C12:1; C14; C14:1; C14:2; C16; C18; C18:1; C18:2; C4OH; C5DC; C5OH; C6DC; C14:OH; C16:OH; C16:1OH; C18OH; C18:1OH; C3DC; C4DC	Markers of fatty acid oxidation and organic acid disorders (i.e., MCAD deficiency)
Amino acids (n=11)	Arginine; phenylalanine; alanine; leucine; ornithine; citruline; tyrosine; glycine; argininosuccinate; methionine; valine	Markers for amino acidopathies (i.e., maple syrup urine disease)
Hemoglobin ratio (n=1)	Fetal hemoglobin (HbF) / Fetal hemoglobin (HbF) + Adult hemoglobin (HbA)	Marker for hemoglobinopathies (i.e., sickle cell disease)
Endocrine markers (n=2)	17-hydroxyprogesterone (OHP), Thyroid Stimulating Hormone (TSH)	Markers for endocrine diseases, specifically congenital hypothyroidism and congenital adrenal hyperplasia
Enzyme markers (n=3)	Biotinidase (BIOT); Galactose-1-Phosphate Uridyltransferase (GALT); immunoreactive trypsinogen (IRT)	Markers for enzymatic activity deficiencies (i.e., biotinidase)

The BORN-Niday Perinatal database, a database maintained by the BORN Ontario provincial birth registry, captures data on all hospital births ≥ 500 grams or with a gestational length ≥ 20 weeks in Ontario. The database includes information on maternal demographics and health behaviours, clinical indications for induction and cesarean delivery, obstetric complications during the pregnancy, type of labour, mode of delivery, and gestational age at birth, which were used to identify term births and preterm births as well as the mode of delivery (Objective 1) and clinical subtype of preterm birth (Objective 2). BORN Ontario collects data from medical records, clinical forms, and interviews with

patients at the time of admission to hospital to give birth. To assess and maintain the quality of the database, periodic quality assurance projects are undertaken to ensure the reliability, completeness and comprehensiveness of the data.⁸²

3.2.1 Data Linkage, Data Access, and Privacy Considerations

This study was conducted at ICES within their secure research environment, adhering to the strict policies and procedures of NSO, BORN and ICES. ICES is a Prescribed Entity under provincial privacy legislation, meaning health information custodians can disclose personal health information about their patients to ICES without express consent. ICES is a repository for health administrative and other individual-level databases. Most databases transferred to ICES have direct personal identifiers (usually health card numbers), which are de-identified and replaced by an encrypted person identifier, known as an ICES Key Number (IKN). Each person in Ontario is assigned their own unique IKN which is necessary for linkage across datasets within ICES. As part of a larger, ongoing program of research, probabilistic and deterministic linkages were employed by ICES data analysts to link the BORN-Niday and NSO datasets with other health administrative databases at ICES. For this project, both the BORN-Niday and NSO data containing the unique IKN were deterministically linked without access to any identifiable personal health information.

3.3 Study Population

First, as part of the ongoing provincial study from which my two studies were an extension, the following exclusion criteria were applied to all births (term and preterm) to derive the source population: birth year prior to 2008, unsatisfactory newborn screening samples (e.g., not enough blood on the sample, missing demographic information), multiple births, duplicate records, birth weight less

than 500 grams, and blood transfusion. Additional objective-specific inclusions and exclusions are described below.

Objective 1: For our study population for Objective 1, we included all infants born in Ontario between January 1, 2010 to March 31, 2012 term gestation (see Figure 4.1). Prior to 2010, fetal-to-adult Hb level was not collected. Therefore, rather than excluding this analyte from our analyses, a restriction of the study period was instead applied. Information recorded in the birth registry was used to measure term births, along with the mode of delivery. Term birth is defined as a live birth at ≥ 37 completed weeks of gestation. The term births were limited to low-risk women with no obstetrical (e.g., gestational diabetes and eclampsia) or intrapartum (e.g. cord prolapse or intrapartum bleeding) complications, or pre-existing maternal (e.g., chronic hypertension or heart disease) health problems (see full list of exclusions within these categories, Table 3.2). This project focused on uncomplicated births to low-risk women so that any associations between screening analytes and the mode by which the infant was delivered would not be confounded by any clinical indication for a cesarean mode of delivery. We selected these criteria based on previous Canadian research^{83,84} and had them reviewed by a high-risk obstetrician at The Ottawa Hospital with extensive subject matter expertise (Oral communication: M Walker, MD; February 28 2018). Liu et al defined low-risk women as those without a pre-specified set of risk factors or obstetric complications, which were similar to those used in our study (See Table 3.2). Using the BORN-Niday database, another study included women with a singleton live birth and with no fetal or maternal health conditions or obstetrical complications to obtain low-risk women.⁸⁴

We classified mode of delivery as: (i) vaginal delivery following spontaneous onset of labour, or (ii) planned cesarean delivery prior to the onset of labour, indicated for three specific clinical scenarios:

breech presentation, repeat cesarean delivery, or maternal request for cesarean. The studies used to define low-risk women were also used to identify these three indications for cesarean.^{83,84} Using the Discharge Abstract Database, Liu et al. defined planned cesarean delivery using breech presentation with no labour as a surrogate because no code was available in the database for an elective cesarean performed on-demand.⁸³ They reasoned that this would minimize the confounding by indication that could severely bias the comparison of modes of delivery because adverse health outcomes were significantly lower among infants with planned cesarean (indicated for breech presentation) compared to vaginal deliveries.^{85,86} Using the same database as this study, Dunn et al. defined planned cesarean delivery using similar indications.⁸⁴

Table 3.2. List of intrapartum and obstetrical complications and maternal health problems leading to exclusion of infants born at term that were not considered to be delivered from low-risk women

<p>Intrapartum complications: Cord prolapse, uterine rupture/dehiscence, intrapartum bleeding, meconium, non-reassuring fetal status, non-progressive labour/lack of descent/dystocia, post-partum hemorrhage, shoulder dystocia, suspected chorioamnionitis, suspected sepsis (unexplained fever)</p> <p>Obstetrical complications: Eclampsia, premature rupture of membranes, preterm labour, premature rupture of membranes, urinary tract infection, other cervical/vaginal infection, gestational diabetes, hypertension (gestational or transient), intrauterine growth restriction/ small for gestational age, large for gestational age, periodontal infection, placenta previa, placental abruption</p> <p>Maternal health problems: Heart disease, hepatitis B, HIV, lupus, asthma, thyroid disease, chronic hypertension, diabetes insulin dependant, diabetes non-insulin dependant, alcohol dependence syndrome, drug & medication use (methadone treatment, narcotics, herbal remedies, opioids, prescription drugs, cocaine, gas/glue sniffing, hallucinogens, marijuana), psychiatric disorders (anxiety</p>

or depression during this pregnancy, previous history of anxiety, postpartum depression, other mental illness)

¹ Given that Dunn et al. used the same data source as the present study, the intrapartum, obstetrical and maternal or fetal health problems exclusions were the same as the present study.

² Medical risk factors and obstetrical complications exclusions made in the Liu et al. study that weren't similar in the thesis project: cerebral hemorrhage, antepartum hemorrhage, liver, renal or thyroid abnormalities, herpes, unengaged head, soft-tissue disorder, uterine scar unrelated to cesarean delivery, congenital fetal central nervous system anomaly, chromosomal abnormality, isoimmunization, oligohydramnios or polyhydramnios, antepartum venous complication, pulmonary embolism.

Objective 2: For our study population for Objective 2, we included all infants born in Ontario between January 1, 2008 to March 31, 2012 at preterm gestation (see Figure 4.9 for study flow diagram).

Characterization of preterm birth subtypes was derived from the following variables recorded in the birth registry: gestational age at birth (completed weeks), labour type, and obstetrical complications [preterm premature rupture of membranes (PPROM)]. Using the gestational age at birth variable, preterm births were restricted to live births at <37 weeks of completed gestation and then further classified into early-preterm (<34 weeks) and late-preterm (34-36 weeks) for closer control of possible confounding by gestational age, as it has been determined to be strongly associated with some newborn screening analytes.¹² These two categories of preterm birth were categorized as such because 34 weeks gestational age is an important clinical threshold as it represents the lower limit of the late-preterm period.⁸⁷ Given sample size limitations for early-preterm infants, we ultimately limited our analyses to late-preterm infants. Following an approach proposed by Savitz et al, we defined the subtypes of preterm birth based on clinical presentation, as 1) spontaneous onset of labour; 2) PPRM and 3) medically-indicated.¹⁸ Firstly, if the record had a code for spontaneous onset of labour and there was no documented PPRM, it was characterized as spontaneous onset of labour prior to membrane rupture. Remaining records with documented PPRM were characterized as preterm premature rupture of membranes. And finally, all other records were characterized as medically-indicated preterm births. This latter group was excluded from our analyses as the medical indication for intervening and delivering an

infant at a preterm gestation would potentially confound any associations between screening analytes and preterm birth subtype.

3.4 Data Cleaning and Preparation

Missing Data

We excluded infants with missing information on gestational age, since this was critical for defining the study population. We also excluded records with missing birth weight since we used this variable in combination with gestational age to identify records with implausible gestational age and birth weight combinations, so they could be excluded from the study population. We also excluded infants missing analyte values since these were the main variables of interest in this project. We opted not to use any imputation methods because only approximately 1% of infants had missing information for any of the analytes and we did not anticipate that imputing these missing values would have a large impact on the results (see Figures 4.1 and 4.9 for study flow diagrams for Objective 1 and Objective 2, respectively).

Specific to Objective 2, approximately half of the infants had missing information for administration of total parenteral nutrition (TPN). In the earlier days of data collection at NSO, it was common practice that missing values signified that no TPN was given. Therefore, we assumed that missing values were indicative of “No TPN” for this objective. In addition, infants with missing information for mode of delivery were excluded.

Outliers

To minimize the impact of extreme analyte values, winsorization of all the analytes was completed. This is a process in which outliers below the 0.001st percentile or above the 99.999th percentile were replaced with the analyte values at those percentile levels.⁸⁸

Standardization of analytes

We standardized all analytes to the standard normal distribution, meaning they all had a mean of 0 and a standard deviation of 1. This was completed so that each analyte would be comparable across a standard measurement scale (units of standard deviation) since raw analytes have different units of measurement. Outlier detection and standardization of analyte values was conducted in the full source population, before objective-specific exclusions were applied. Standardization was conducted by calendar week of birth to minimize variation due to external factors, including seasonal variation and laboratory factors (e.g., assay changes). For the analyte ratios, they were created as a ratio first and the same manner in which the individual analytes were standardized was also completed for the ratios.

Implausible Birth weight and Gestational Age Combinations

We computed birth weight z-scores to identify records for which the infant's birth weight was incompatible with the documented gestational age using an algorithm described by Basso and Wilcox.^{89,90} First, the median birth weight within each sex-and gestational-age specific category between 25-42 weeks in our source population was computed. We then used a Canadian birth weight reference standard developed by Kramer et al.⁹¹ to obtain sex-and-gestational-age specific category estimates of birth weight standard deviations (SD). A birth weight z-score was computed for each infant by subtracting the stratum-specific median birth weight from the infant's observed birth weight and dividing by the corresponding stratum-specific SD from the reference standard.

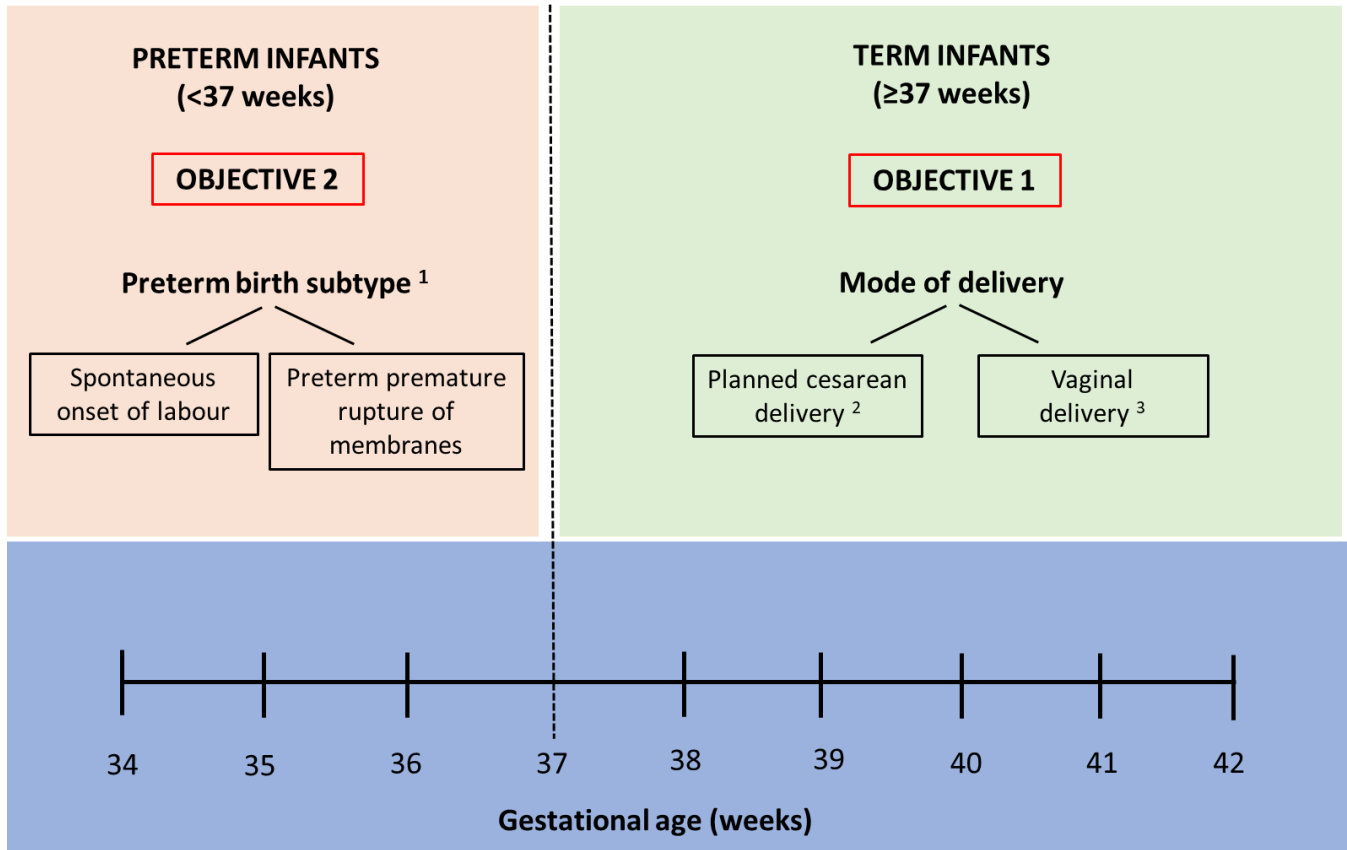
We used different criteria to identify term and preterm infants as having implausible birth weight–gestational age combinations.^{89,90} For births with a documented gestational age ≥ 37 completed weeks, absolute z-scores of 5 or higher were flagged as implausible and excluded. However, for records with a documented gestational age < 37 weeks, we used an asymmetric cut-off since incompatible birth weight–gestational age combinations tend to arise mainly among heavier preterm babies. Thus, for infants with a documented gestational age < 37 weeks, we excluded infants that did not have a z-score between -4 and +3 SD.

3.5 Dependent Variables

Objective 1: Infants were characterized into two categories for mode of delivery: (i) planned cesarean delivery (indicated for breech presentation, previous cesarean, or maternal request) with no labour, or (ii) vaginal delivery following spontaneous onset of labour (see Figure 3.1).

Objective 2: We defined two categories of preterm birth: (i) those following spontaneous onset of labour, or (ii) those following preterm premature rupture of membranes (PPROM) (see Figure 3.1).

Figure 3.1 Graphical presentation of dependent variables for Objectives 1 and 2



¹ The subtypes were categorized according to the initiation of events that ultimately lead to a preterm delivery, irrespective of whether it resulted in a vaginal or cesarean delivery.

² Planned cesarean delivery was defined as a cesarean delivery that has been indicated for breech presentation, previous cesarean section or maternal request prior to the onset of labour.

³ Vaginal delivery was defined as vaginal delivery following spontaneous onset of labour.

3.6 Independent Variables

There were 48 individual analytes available in the NSO database. These analytes included: acyl-carnitines (n=31), amino acids (n=11), fetal-to-adult hemoglobin level (n=1), endocrine markers (n=2), and enzyme markers (n=3) (See Table 3.1).

Objective 1: We included all 48 individual analytes in our analyses. Additionally, rather than using a traditional approach which would require testing a number of pairwise interactions between individual

analytes, we chose to approximate interaction terms by computing analyte ratio combinations (n=1,128) which allows us to capture relative levels of analytes. We also included a number of clinical covariates in our models, including infant sex (male vs. female), continuous gestational age (weeks), continuous birth weight (grams), mode of feeding (TPN vs. other), and age at blood spot collection (hours).

Objective 2: Given that the study period for this objective is between 2008 to 2012 and that fetal-to-adult hemoglobin levels were not collected until 2010, we decided to exclude this analyte rather than further restrict our study period to ensure a large enough sample size and statistical power for modelling. Therefore, a total of 47 individual analytes were included in the analyses for this objective, as well as analyte ratio combinations (n=1, 081). We also included the same clinical covariates as Objective 1, with the exception of TPN. Approximately half of the infants had missing information for TPN and we did not think this justified the exclusion of half our study population in our analysis since TPN occurs only in a small portion of the population. In the earlier days of data collection at NSO, it was common practice that missing values signified that no TPN was given. Therefore, we assumed the missing values indicated no TPN and assessed its association in a sensitivity analysis. Finally, as determined in Objective 1, quantitative values of analytes do differ by mode of delivery, therefore, this variable was also included as a covariate for Objective 2.

3.7 Statistical Analyses

Descriptive Analyses

Extensive descriptive statistical analyses were conducted where baseline characteristics of the study population were assessed. Statistical graphics, including correlation heat maps, were generated to investigate the relationships between analytes, and to evaluate differences in individual analyte levels or

correlation patterns between the modes of delivery for low-risk term births (Objective 1) and preterm subtypes for preterm births (Objective 2).

Previous research has reported that newborn screening analyte levels were highly correlated with gestational age,¹² therefore, we produced histograms (Appendix C and D) and boxplots (Appendix E and F) of each analyte stratified by gestational age in weeks to determine if values qualitatively differed sufficiently to warrant gestational-age specific models. After concluding that gestational age-specific models were not required, we opted to address confounding by gestational age by including adjustment in each of our models. For Objective 1, we further assessed gestational age in sensitivity analyses where we stratified our model into early term (37-39 weeks) and full-term (≥ 40 weeks). For Objective 2, gestational age was further controlled for by restricting our analyses to late-preterm infants (34-36 weeks).

Model Building

To ensure the stability of the models, we determined a priori that the maximum number of variables that could be included in our models was based on a requirement of 10 cases of the outcome (Objective 1: planned cesarean delivery; Objective 2: PPRM) per model parameter.⁹² Prior to modelling, we used several steps to reduce the number of analyte ratios. We computed partial Spearman correlations between cesarean delivery or PPRM and each individual standardized analyte ratio. We then prioritized the analyte ratios that were most correlated with the outcome, specifically the top 100 ranked partial Spearman correlation analyte ratios for inclusion in our model, in addition to the individual analytes and clinical variables. Spearman correlation is a statistic used to determine the strength and direction of association between two variables, and, unlike Pearson correlation, it is

robust to both departures from linearity and outliers.^{93,94} We then computed partial Spearman correlations (i.e., adjusted for the covariates identified in Section 3.6) between cesarean delivery or PPRM and individual analytes as well as the standardized analyte ratios. Additionally, Spearman correlation coefficients can be interpreted in both directions.

A multivariable logistic regression model was implemented to investigate the association between newborn screening analytes and mode of delivery among low-risk infants born at term (Objective 1) and preterm birth subtypes among preterm infants (Objective 2). For both objectives, we began with a pre-specified base model (Model 1) including the following independent variables: infant sex (male versus female), gestational age at birth (weeks), timing of blood collection (hours), mode of feeding (total parenteral nutrition versus other), and birth weight (grams) to demonstrate the initial performance of the model with clinical variables only. For Objective 2, mode of delivery (vaginal delivery versus cesarean delivery) was also included as a covariate because results from Objective 1 of this thesis demonstrated that this covariate was associated with newborn screening analytes. Additionally, for Objective 2, total parenteral nutrition (TPN) was not included (refer to Section 3.6 for rationale on exclusion of this variable in the main analysis). Afterwards, a second model (Model 2) was built and included solely the individual analytes as well as the top 100 ranked analyte ratios based on partial Spearman correlations as variables of interest. A final model (Model 3) was built and included all the clinical variables and the analytes and the top 100 ranked analyte ratios (i.e., Model 1 + Model 2 variables) to examine the performance of the model with the inclusion of newborn screening analytes data.

We employed an automated backward stepwise selection approach by initially fitting each model (Model 1, Model 2, and Model 3) with all variables of interest described above and then variables not

meeting pre-specified criteria were dropped.⁹⁵ A significance level of 0.1 was required to allow a variable to stay in the model. The process of re-fitting the reduced models was continued until none of the remaining covariates met the criterion for removal from the model. Backward selection was chosen over forward selection because each addition of a new covariate to the model can render the already included variables non-significant.⁹⁶ Odds ratios (OR) with 95% confidence intervals (CI) were computed. We ranked the effect size from highest to lowest based on absolute values of corresponding beta coefficients (log odds ratio) to facilitate our assessment of which analytes were most strongly associated with the mode of delivery and preterm birth subtypes. All statistical analyses were performed using SAS statistical software version 9.4 (SAS Institute, Inc.) and with R/R Studio Desktop version 0.98.1091 using the RMS and HMISC packages.⁹⁷

Test for Linearity

Objective 1: We decided to assess whether using restricted cubic spline terms for the analyte ratios in Model 3 to account for non-linear associations would change the model performance. We selected the top five analyte ratios that were more associated with planned cesarean delivery and assessed the benefit of splining these variables in a model. Based on the likelihood ratio test, we found that incorporation of the splines for the top 5 analyte ratios as cubic spline terms demonstrated a statistical improvement in fit. However, this did not result in any meaningful differences in the interpretation of the full multivariable model, and therefore, we opted not to include the cubic spline terms in our final models.

Objective 2: We adopted the same approach to test for linearity as we did in Objective 1. However, in this case, when we modelled the top five analyte ratios as cubic spline terms, the fit of the model was

not significantly improved as demonstrated by a likelihood ratio test, and therefore, cubic spline terms were not included in the final model.

Model Performance Assessment

There are two key elements for assessing the performance of a regression model: model calibration and model discrimination.⁹⁸ Model discrimination refers to the ability of a model to discriminate between those with and without the outcome and is often measured by the concordance statistic (c-statistic).⁹⁸ The c-statistic is equivalent/analogous to the area under the curve (AUC). This curve is known as the receiver operating characteristic (ROC) and is a plot of the sensitivity against the false-positive rate (1-specificity). We used the c-statistic to assess discriminative ability of the models. A c-statistic value of 0.5 indicates that the model is no better than chance.⁹⁹ Models with a c-statistic greater than 0.6 and greater than 0.8 are considered to exhibit reasonable and strong performance, respectively.⁹⁹

Model calibration refers to the agreement between observed outcomes and predicted outcomes.

Although the Hosmer-Lemeshow test can be used to assess calibration/fit for logistic models, it is overly powerful in large samples, often indicating a statistically significant lack of fit, where the actual lack of fit is minor and not of practical/clinical importance.¹⁰⁰ Therefore, we assessed calibration by constructing and visually examining calibration plots. We divided our study population into deciles based on predicted probability of cesarean delivery (Objective 1) or PPRM (Objective 2). We then compared the mean predicted probabilities of cesarean delivery or PPRM with the mean observed probability of the outcome within the deciles. Log transformations were completed prior to plotting because the deciles were strongly left skewed. When a model is well calibrated, the mean predicted probabilities and mean observed probabilities should be close in value. Graphically, this scenario would be represented by data points that lie close to the 45-degree line through the origin (also known as the reference line).

Model Validation

Model validation is also an important aspect in assessing model performance. However, external validation of the models was not performed in the current study because it was not possible to obtain a new sample of the same population or of a sample from a similar population. Therefore, internal validation techniques were employed.

Objective 1: Given the large sample size of infants born at term gestation to low-risk women, we were able to partition the data into two datasets: a training dataset and a validation dataset. These subsamples were generated by randomly partitioning records according to a 2/3 to 1/3 ratio for training and validation datasets, respectively. The purpose of partitioning the data is to fit the model in the training dataset and then assess the performance of the model in the independent validation dataset. Using the same sample for fitting and evaluation of performance can potentially lead to over-fitting and overestimation of the performance metrics of the model.¹⁰¹

Objective 2: Due to the smaller sample size, data partitioning was not practical for Objective 2. Instead, we conducted an internal validation using a cross-validation approach. Cross-validated predicted probabilities were derived from the leave-one-out cross-validation (LOOCV) principle to provide ROC analysis.¹⁰² The LOOCV approach uses n-1 participants for model development and model validation on the participant who was left out. Therefore, there are as many training and validation splits as there are observations.¹⁰² By validating our model using cross-validation, we were able to compute performance metrics of the model that would be less susceptible to overfitting and better reflect how the model might perform in new data.

3.8 Sensitivity Analyses

The same approach of ranking the effect size based on corresponding beta coefficients (log odds ratio) as was completed in the original analyses, was adopted for the sensitivity analyses. A comparison of the analytes with the highest ranked and lowest ranked ORs were compared to see if the same analytes were strongly associated with mode of delivery (Objective 1) or preterm birth subtypes (Objective 2) in both original and sensitivity analyses.

Objective 1: We performed three sets of sensitivity analyses. For the first sensitivity analyses, we stratified the term infants into early-term (37-39 weeks) and full-term (≥ 40 weeks) infants to further control for gestational age as a confounder. Additionally, given that the timing of blood spot collection was different between infants with planned cesarean delivery and vaginal delivery following spontaneous onset of labour, we restricted our second sensitivity analyses to infants who obtained a blood spot collection within the recommended window of 24-48 hours. Finally, given that we only included the top 100 ranked partial Spearman ratios, we still had degrees of freedom to accommodate additional parameters in the model. Thus, as our last sensitivity analysis, we added the remaining analyte ratios according to rank order based on partial Spearman's correlations until the maximum number of parameters was reached.

Objective 2: Four sets of sensitivity analyses were performed for Objective 2. Given that we excluded infants who underwent a transfusion from the original analyses, we decided to assess the impact of including these infants as our first sensitivity analysis because blood transfusions are known to affect the results of galactosemia screens and may affect other analyte values as well. Next, we restricted our analyses to infants born between January 1, 2010 and March 31, 2012 and included the fetal-to-adult

hemoglobin level, as this was collected at NSO starting in 2010. As the third sensitivity analysis, we included TPN in our model as total parenteral nutrition is important with respect to quantitative analyte levels and because we assumed that missing status for TPN indicated no TPN was administered (see Section 3.6 for justification for its exclusion in the original model). Given that we still had degrees of freedom leftover since we only included the top 100 ranked ratios based on partial Spearman correlations, as our last sensitivity analyses, we added the remaining analyte ratios according to rank order based on partial Spearman's correlations until the maximum number of parameters was reached.

3.9 Ethical Considerations

The study received approval from the institutional review board at Sunnybrook Health Sciences Centre/ICES Privacy Office, and the Children's Hospital of Eastern Ontario (See Appendix A for approval forms). In addition, a letter of support from Newborn Screening Ontario and BORN was obtained (See Appendix B). Ottawa Health Science Network – Research Ethics Board (OHSN-REB) approval was not required for this project as they have a memorandum of understanding with Sunnybrook/ICES such that database only studies are reviewed through the ICES centralized approval process in Toronto.

CHAPTER 4. RESULTS

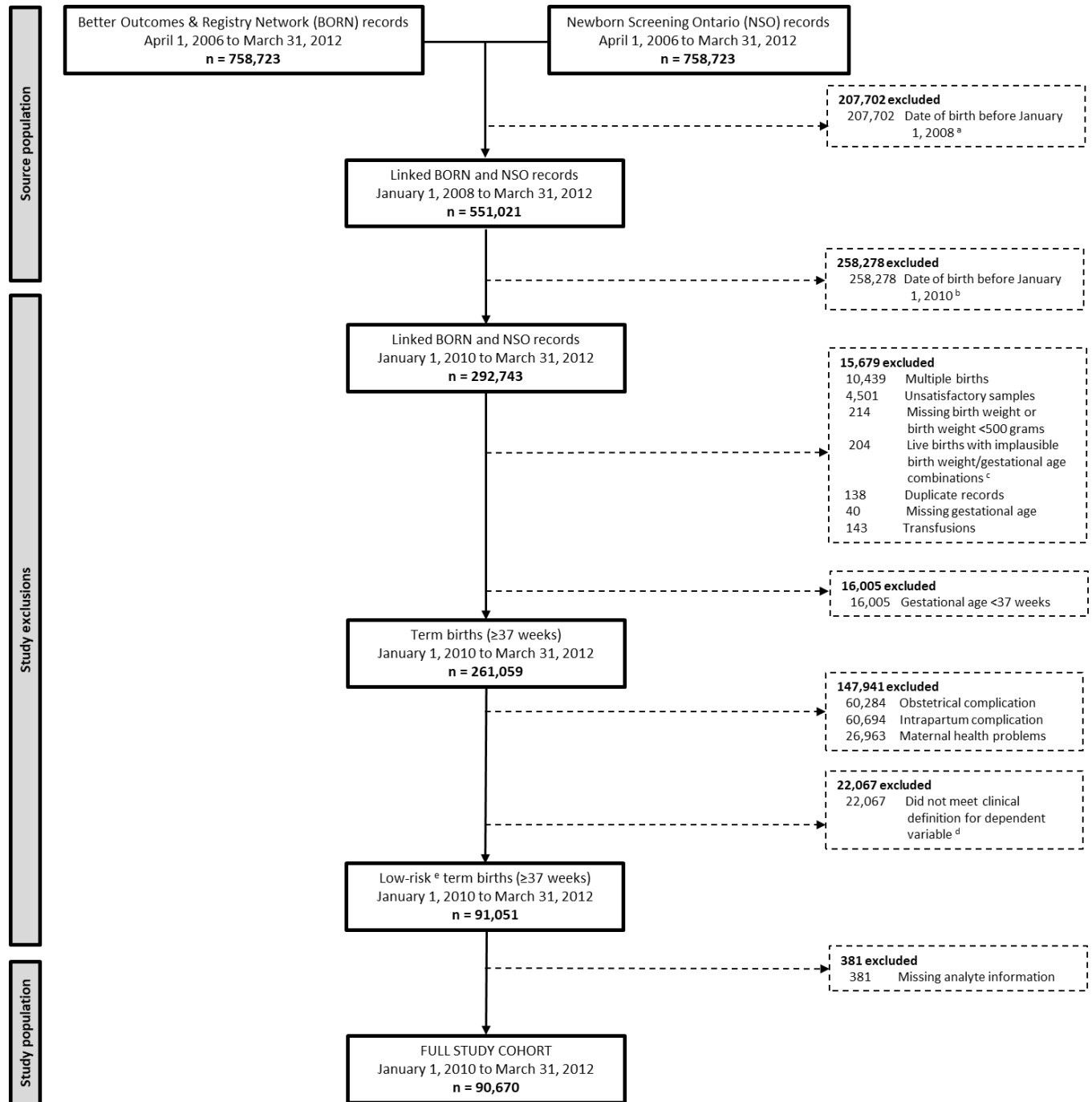
4.1 Objective 1

4.1.1 Characteristics of the Study Population

Following study exclusions, there were 90,670 infants in the full study cohort for Objective 1 (see Figure 4.1 for study flow diagram). In this cohort of low-obstetrical risk deliveries, 72,684 infants (80.2%) were born vaginally following spontaneous onset of labour, and the remaining 17,986 infants (19.8%) were delivered following planned cesarean delivery (with no labour) for indications of breech presentation, repeat cesarean or maternal request. The distribution of characteristics was similar across both modes of delivery with exception of gestational age and age at blood spot collection. Infants delivered by planned cesarean had an earlier blood spot collection at a median of 26.4 hours (IQR 24.5 – 35.1) compared to infants delivered vaginally following spontaneous onset of labour (median: 40.9 hours, IQR 27.3 – 47.9) Approximately half of planned cesarean deliveries occurred at a gestation of 38 weeks, whereas a third of vaginal delivery following spontaneous onset of labour occurred at 40 weeks gestation (Table 4.1).

After randomly partitioning the full study cohort into training (n= 60,477; 66.7%) and validation (n= 30, 193; 33.3%) subsets, there were 48,522 infants (80.2%) born by vaginal delivery following spontaneous onset of labour and the remaining 11,955 infants (19.8%) born by planned cesarean delivery with no labour in the training dataset, and 24,162 infants (80.0%) born by vaginal delivery following spontaneous onset of labour and 6,031 infants (20.0%) born following planned cesarean delivery with no labour in the validation dataset. The distribution of the characteristics was maintained as in the full study cohort (See Table 4.1).

Figure 4.1 Study flow diagram for Objective 1



^a Prior to 2010, fetal-to-adult hemoglobin level was not collected. Therefore, rather than excluding this analyte for analysis, a restriction of the study period was instead applied.

^b The exclusions of births prior to 2008 was applied to derive the source population. Additional objective-specific exclusions were applied afterwards (as demonstrated by the headings on the left).

^c Algorithm similar to that described by Basso and Wilcox⁹⁰ to clean gestational age values on live birth records was used.

^d Was not considered planned cesarean delivery or vaginal delivery following spontaneous onset of labour.

^e Low-risk is defined as live births with no history of pre-existing maternal chronic health problems, and no obstetrical or intrapartum complications

Table 4.1 Baseline characteristics of full, training and validation datasets for singleton term live births in Ontario between 2010 and 2012 by mode of delivery

	Full dataset n = 90,670				Training dataset n = 60,477				Validation dataset n = 30,193			
Characteristics	Mode of delivery				Mode of delivery				Mode of delivery			
	Planned cesarean delivery with no labour		Vaginal delivery following spontaneous onset of labour		Planned cesarean delivery with no labour		Vaginal delivery following spontaneous onset of labour		Planned cesarean delivery with no labour		Vaginal delivery following spontaneous onset of labour	
	n	%	N	%	n	%	n	%	n	%	n	%
All infants	17,986	100	72,684	100	11,955	100	48,522	100	6,031	100	24,162	100
Infant sex												
Female	9,024	50.2	36,341	50.0	5,904	49.4	24,291	50.1	3,058	50.7	12,052	49.9
Male	8,962	49.8	36,343	50.0	6,051	50.6	24,231	49.9	2,973	49.3	12,110	50.1
Gestational age (completed weeks)												
Mean ± SD	39.2 ± 1.0		38.5 ± 0.8		38.5 ± 0.8		39.2 ± 1.0		38.5 ± 0.8		39.2 ± 1.0	
37	999	5.6	4,476	6.2	656	5.5	2,966	6.1	343	5.7	1,510	6.3
38	8,615	47.9	13,025	17.9	5,795	48.5	8,719	18.0	2,820	46.8	4,306	17.8
39	6,913	38.4	25,158	34.6	4,572	38.2	16,836	34.7	2,341	38.8	8,322	34.4
40	1,104	6.1	24,521	33.7	701	5.9	16,357	33.7	403	6.7	8,164	33.8
>40	355	2.0	5,504	7.6	231	1.9	3,644	7.5	124	2.1	1,860	7.7
Birth weight (grams)												
Mean ± SD	3,402.0 ± 422.9		3,410.1 ± 436.5		3,407.1 ± 436.0		3,403.0 ± 423.5		3,415.9 ± 437.3		3,402.2 ± 421.7	
<3000	2,985	16.6	12,220	16.8	2,003	16.8	8,190	16.9	982	16.3	4,030	16.7
3000-3499	7,908	44.0	31,610	43.5	5,279	44.2	21,103	43.5	2,629	43.6	10,507	43.5
3500-3999	5,425	30.1	22,778	31.3	3,581	29.9	15,138	31.2	1,844	30.6	7,640	31.6
≥4,000	1,668	9.3	6,076	8.4	1,092	9.1	4,091	8.4	576	9.6	1,985	8.2
Age at blood spot collection (hours)												
Median (IQR)	26.4 (24.5 – 35.1)		40.9 (27.3 – 47.9)		26.5 (24.5 – 35.1)		40.6 (27.0 – 47.9)		26.4 (24.5-35.0)		41.2 (27.7 – 48.0)	
≤72	17,644	98.1	68,409	94.8	11,735	98.2	45,966	94.7	5,904	97.9	22,942	94.9
>72	341	1.9	3,749	5.2	219	1.8	2,556	5.3	127	2.1	1,220	5.1
Missing	1	0.0	0	0.0	1	0.0	0	0.0	1	0.0	0	0.0
Any total parenteral nutrition ^a												
No	17,777	99.6	71,860	99.6	11,910	99.6	48,324	99.6	6,005	99.6	24,061	99.6
Yes	28	0.2	48	0.1	20	0.2	29	0.1	8	0.1	19	0.1
Missing	43	0.2	250	0.3	25	0.2	169	0.3	18	0.3	82	0.3

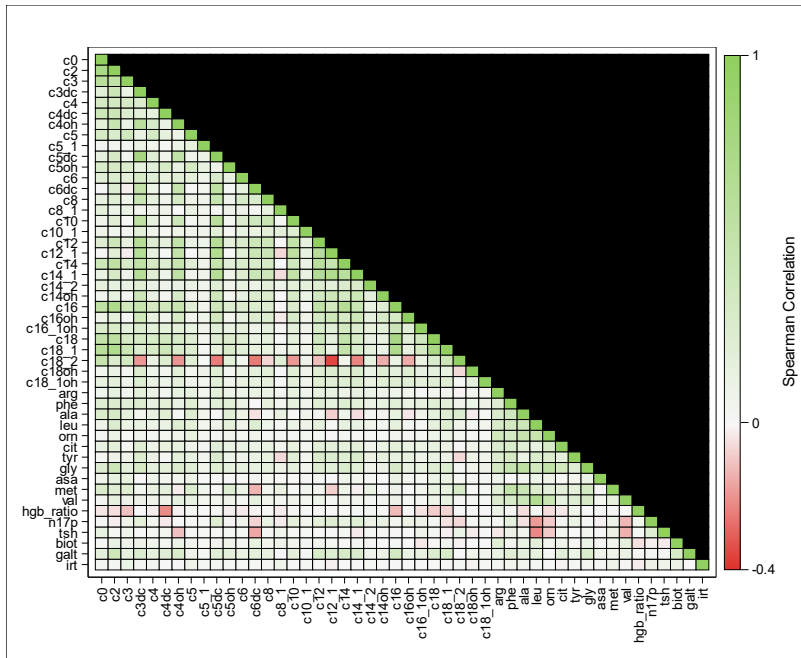
IQR: inter-quartile range; SD: standard deviation

^a Total parenteral nutrition alone or in combination with other infant feeding method.

4.1.2 Heat Map Analysis

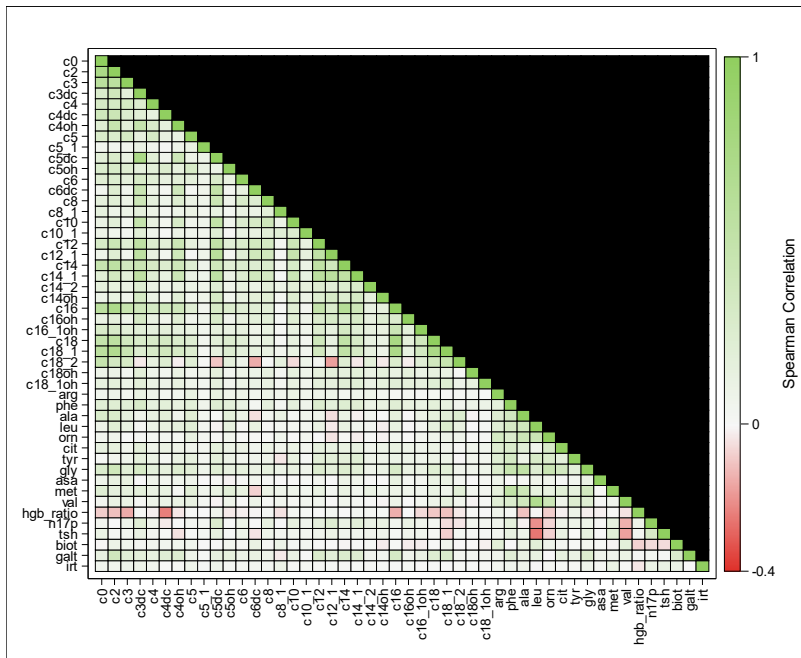
Among infants with planned cesarean delivery, the strongest positive correlation was between malonylcarnitine (C-3DC) and glutarylcarnitine (C-5DC) ($\rho= 0.785$), while the strongest negative correlation was dodecanoylcarnitine (C-12:1) and octadecadienoylcarnitine (C-18:2) ($\rho= - 0.351$) (See Figure 4.2). Among those with vaginal delivery following spontaneous onset of labour, the strongest positive correlation was between hexadecanoylcarnitine (C-16) and octadecenoylcarnitine (C-18:1) ($\rho= 0.763$), while the strongest negative correlation was between thyroid stimulating hormone (TSH) and leucine ($\rho= - 0.253$) (Figure 4.3). Refer to Table 3.1 for a description of the newborn screening analytes.

Figure 4.2 Correlation heatmap of analytes for infants with planned cesarean delivery



Darkener green areas indicate higher levels of positive correlation between analytes, and darker red areas indicate higher levels of negative correlation between analytes.

Figure 4.3 Correlation heatmap of analytes for infants with vaginal delivery following spontaneous onset of labour



Darkener green areas indicate higher levels of positive correlation between analytes and darker red areas indicate higher levels of negative correlation between analytes.

4.1.3 Partial Correlation Analysis

The five strongest partial Spearman correlations between analytes or analyte ratios and planned cesarean delivery were glycine, leucine : valine, malonylcarnitine (C-3DC) : alanine, methylglutarylcarnitine (C-6DC), and leucine. The full results of the partial correlation analysis are presented in Appendix G.

4.1.4 Overall Model Performance

Discrimination

Model discrimination was measured by the c-statistic (see Table 4.2 for results). In the baseline model (Model 1) containing only clinical variables, the c-statistic obtained from the validation dataset was 0.727. In Model 2 in which we included only the 48 individual analytes and the top 100 ranked partial Spearman analyte ratios, the validated c-statistic was 0.784. With the inclusion of clinical variables and the analytes and analyte ratios, the validated c-statistic for Model 3 increased to 0.812. Model 3 performed the best compared to the clinical variables-only model and analytes/analyte ratios-only model.

Table 4.2 C-statistics for discrimination performance for Objective 1

Model		c-statistic (Training dataset)	c-statistic (Validation dataset)
Model 1 ^a	Total parenteral nutrition, age at collection, sex, gestational age, birth weight ^d	0.730	0.727
Model 2 ^b	48 individual analytes and 100 top ranked partial Spearman correlations for analyte ratios ^d	0.782	0.784
Model 3 ^c	Total parenteral nutrition, age at collection, sex, gestational age, birth weight, 48 individual analytes and 100 top ranked partial Spearman correlations for analyte ratios ^d	0.811	0.812

^a Model including the clinical variables only

^b Model including the analytes and analyte ratios only

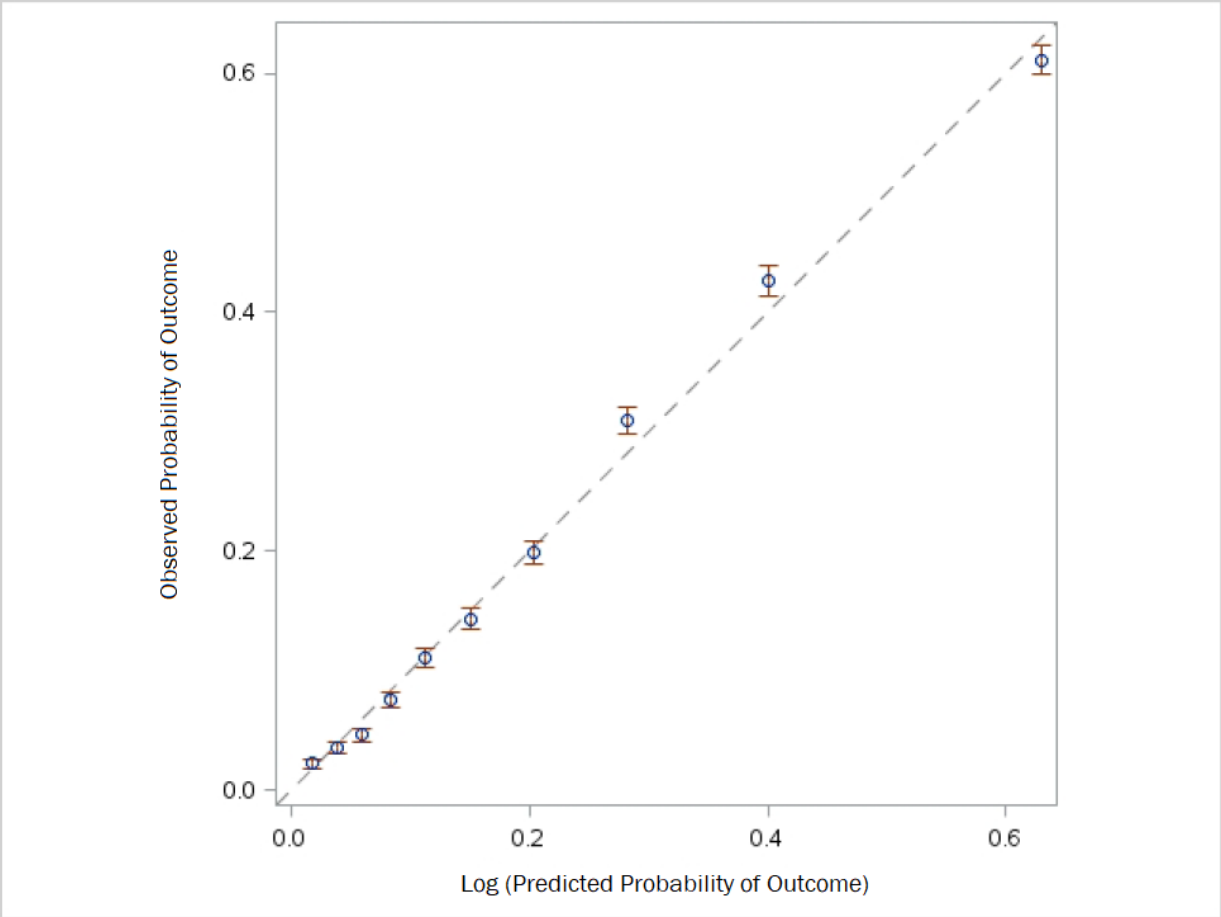
^c Model with a combination of the clinical variables and analytes and analyte ratios

^d These variables were covariates that were entered into the model, not necessarily the covariates that were retained following automated backward selection

Calibration

Figure 4.4 illustrates the calibration plot for term infants comparing the observed probability of planned cesarean delivery across deciles of predicted probability from the model (Model 3). The overall calibration of the model was good but there were several instances where the model over or underpredicted the outcome.

Figure 4.4 Decile calibration plot for term infants



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Association of newborn screening analytes with mode of delivery

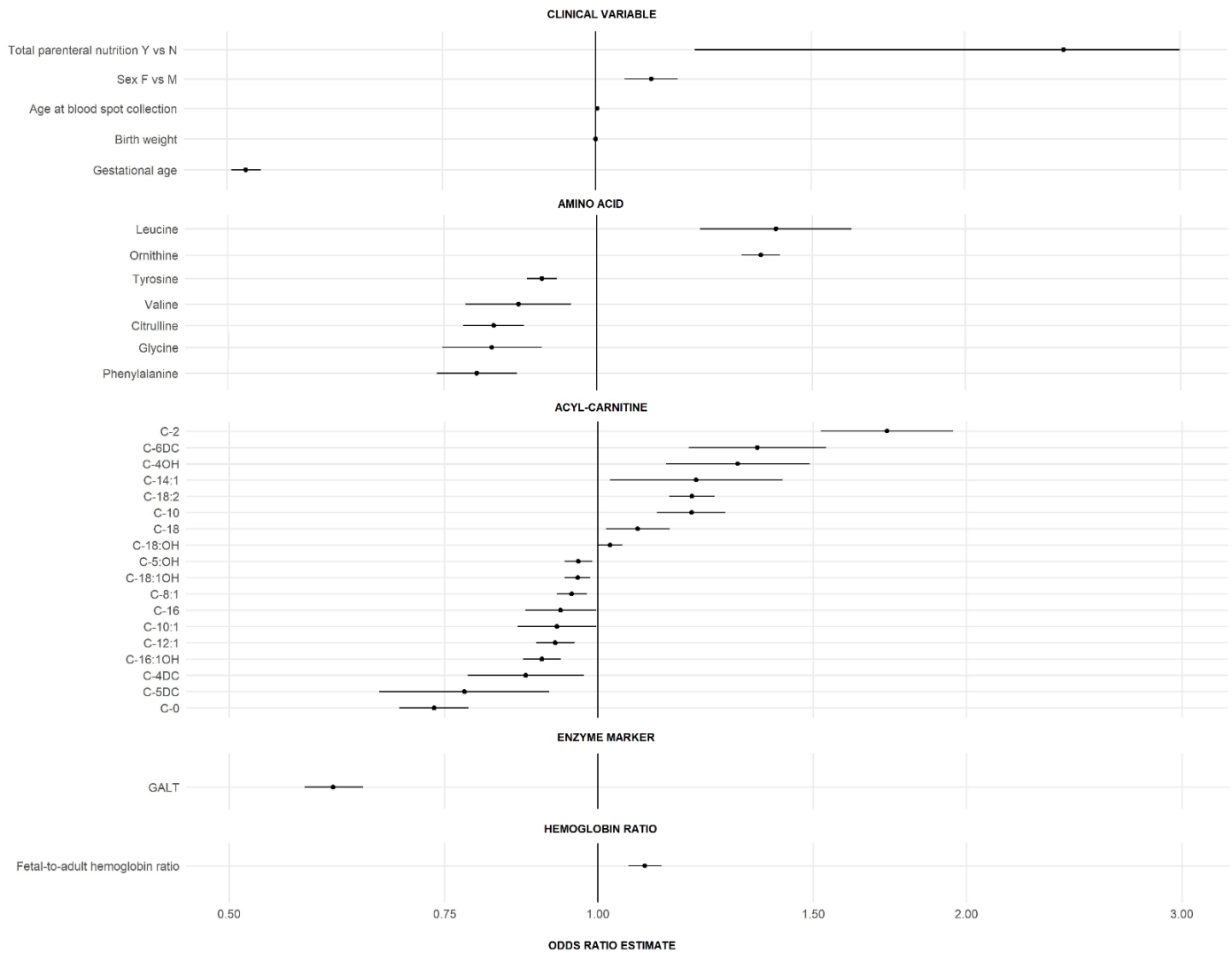
The screening analytes that had the strongest positive associations with planned cesarean delivery were amino acids and acyl-carnitines individually, as well as within several ratios that included amino acids and acyl-carnitines. Specifically, several ratios including leucine, glycine and shorter chain acyl-carnitines were most strongly positively associated with planned cesarean delivery. As compared with vaginal delivery following spontaneous onset of labour, a 1 standard deviation increase in each of the following analytes was associated with an increase in the odds of planned cesarean delivery: C-2 (OR: 1.72; 95% CI: 1.52-1.95), leucine : valine (OR: 1.51; 95% CI: 1.40-1.64), leucine (OR: 1.40; 95% CI: 1.22-1.62), C-3DC : alanine (OR: 1.39; 95% CI: 1.14-1.69), ornithine (OR: 1.36; 95% CI: 1.31-1.41), C-6DC (OR: 1.35; 95% CI: 1.19-1.54), C-4OH (OR: 1.30; 95% CI: 1.14-1.40), C-14 : glycine (OR: 1.29; 95% CI: 1.08-1.53), C-5DC : glycine (OR: 1.27; 95% CI: 1.04-1.55), and C-5DC : citrulline (OR: 1.26; 95% CI: 1.14-1.40). These analytes were the top 10 analytes that were positively associated with planned cesarean delivery.

In contrast, ratios including alanine and citrulline were significantly inversely associated with planned cesarean delivery, compared with vaginal delivery. There were also several individual enzyme and endocrine markers, as well as ratios, that were inversely associated with planned cesarean delivery, specifically those including GALT and TSH (See Figure 4.5a and Figure 4.5b). As compared with vaginal delivery following spontaneous onset of labour, a 1 standard deviation increase in each of the following analytes was associated with a decrease in the odds of planned cesarean delivery: leucine : glycine (OR: 0.44; 95% CI: 0.39-0.51), GALT (OR: 0.61; 95% CI: 0.58-0.64), C-6DC : valine (OR: 0.67; 95% CI: 0.61-0.74), alanine : methionine (OR: 0.72; 95% CI: 0.69-0.74), TSH : biotinidase (OR: 0.73; 95% CI: 0.71-0.75), C-0 (OR: 0.74; 95% CI: 0.69-0.78), C-2 : alanine (OR:

0.76; 95% CI: 0.68-0.85), C-5DC (OR: 0.78; 96% CI: 0.66-0.91), alanine : valine (OR: 0.78; 95% CI: 0.68-0.85), and C-4:OH : citrulline (OR: 0.78; 95% CI: 0.71-0.86). These analytes were the top 10 analytes that were inversely associated with planned cesarean delivery.

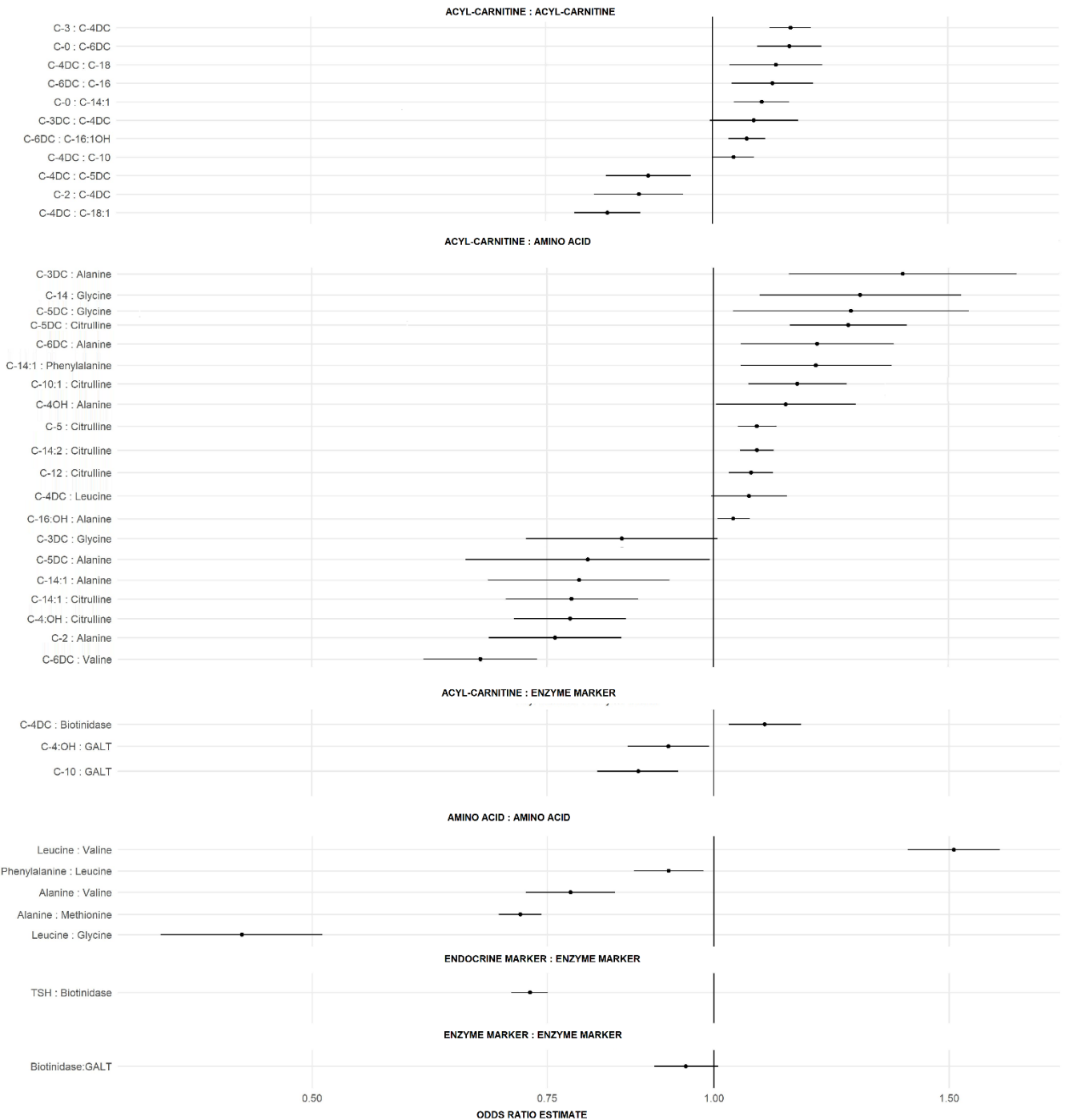
Several clinical variables were also positively associated with planned cesarean delivery. Infants who received total parenteral nutrition (TPN) had an increased odds of 2.41 (95% CI: 1.21-4.83) of planned cesarean delivery, compared to infants who did not receive TPN. Female infants had an increased odds of 1.11 (95% CI: 1.06-1.17) of planned cesarean delivery, compared to male infants. Conversely, gestational age was significantly inversely associated with planned cesarean delivery (OR: 0.52; 95% CI: 0.50-0.53). The remaining clinical variables including birth weight and age at blood spot collection both had odds ratios of 1.00 (95% CI: 1.00-1.00), indicating no association with planned cesarean delivery.

Figure 4.5a Individual analytes and clinical variables in the final model for term infants



Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Figure 4.5b Analyte ratios in the final model for term infants



Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

4.1.5 Overall Model Performance from Sensitivity Analyses

Discrimination

When restricting to early-term (37-39 weeks) infants and infants who had a blood spot collection within the recommended window of 24-48 hours after birth, the results were comparable to the original analyses, with validated c-statistics of 0.794 and 0.792, respectively. However, for the full-term (≥ 40 weeks) infants, the validated c-statistic decreased to 0.683, suggesting substantial overfitting may have occurred in the training data. After inclusion of remaining ranked analyte ratios until the maximum number of parameters was reached, the validated c-statistic increased to 0.814. The results from the sensitivity analyses are summarized in Table 4.3.

Table 4.3 C-statistics for sensitivity analyses for Objective 1

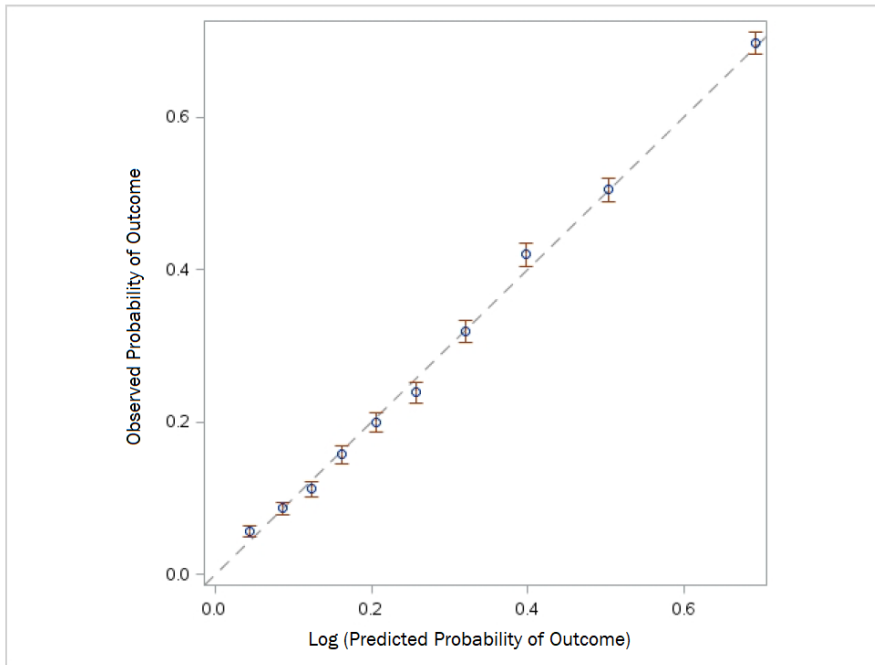
Model	C-statistic (Training dataset) ^a	C-statistic (Validation dataset) ^a
Original analyses	0.811	0.812
Early-term infants (37-39 weeks)	0.771	0.794
Full-term infants (≥ 40 weeks)	0.750	0.683
Infants who had blood spot collection within recommended window	0.831	0.792
Inclusion of remaining ranked analyte ratios until the maximum number of parameters was reached	0.826	0.814

^a The c-statistic shown was obtained from the model with the combination of clinical variables and individual analytes and analyte ratios

Calibration

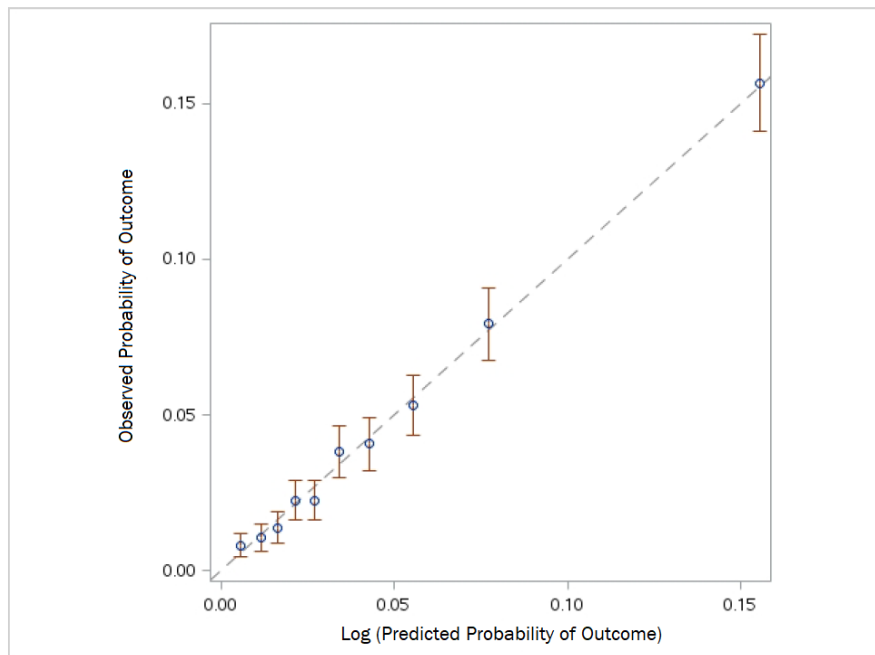
All four models from the sensitivity analyses exhibited similar patterns for calibration (see Figure 4.6 - Figure 4.9). The observed probability of planned cesarean delivery for the first decile consistently exceeded the predicted probability of cesarean delivery across all sensitivity analyses. Calibration improved with increasing decile of predicted probability, agreeing closely with observed probability.

Figure 4.6 Decile calibration plots for early-term infants



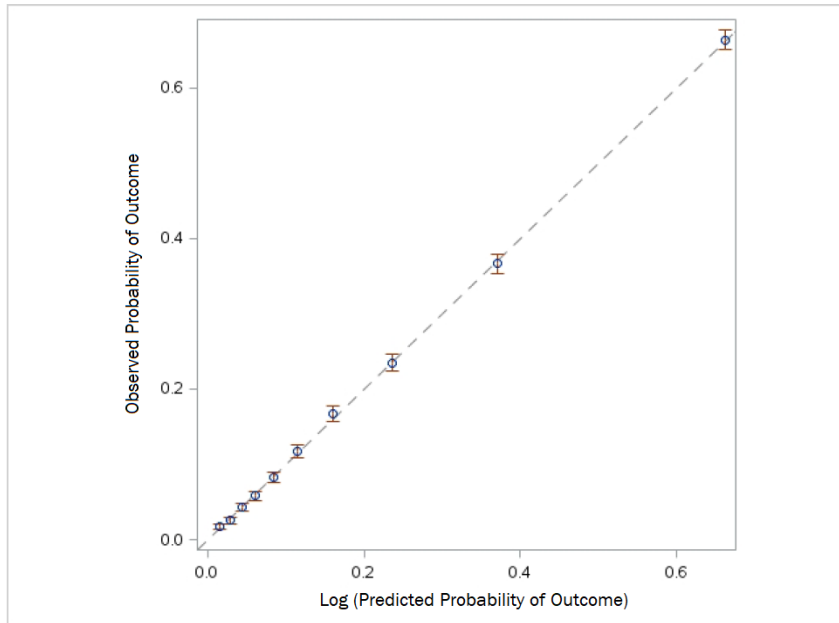
The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.7 Decile calibration plots for full-term infants (≥ 40 weeks)



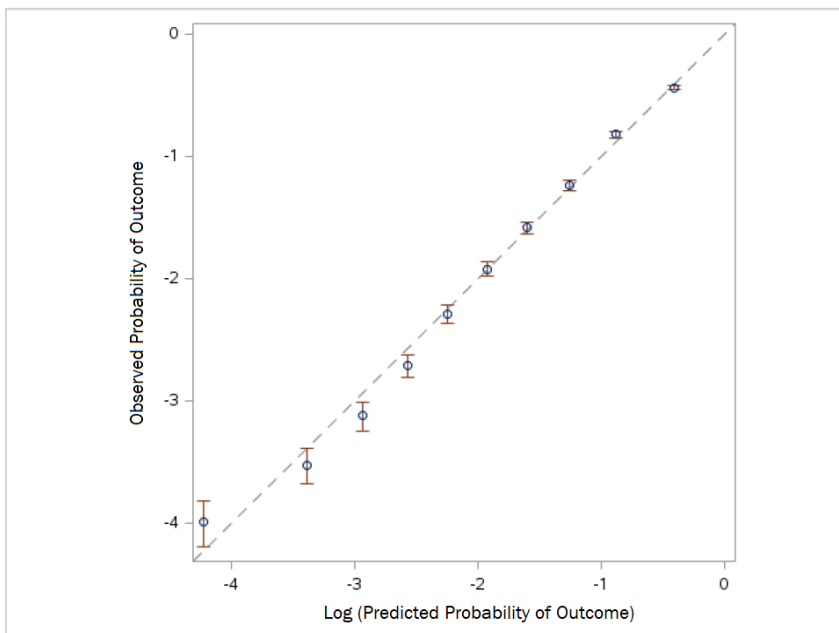
The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.8 Decile calibration plots for infants who had blood spot collection within recommended window



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.9 Decile calibration plots after inclusion of remaining ranked analyte ratios until the maximum number of parameters was reached



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Association of newborn screening analytes with mode of delivery from sensitivity analyses

Many of the findings relating to analytes that were positively and inversely associated with planned cesarean delivery remained similar across all sensitivity analyses with slight variations. For instance, as seen in the main analysis, many shorter chain acyl-carnitines and amino acids and ratios including these analytes were positively associated with planned cesarean delivery, while few to no enzyme or endocrine markers or ratios of these analytes were observed to be positively associated with planned cesarean delivery. Results that were notably different were observed in the sensitivity analysis restricting to full-term infants only (≥ 40 weeks), in which a 1 standard deviation increase in the ratio of citrulline and fetal-to-adult Hb level was associated with a 2-fold increased odds (OR: 2.08, 95% CI: 1.33-3.25) of planned cesarean delivery compared to vaginal delivery following spontaneous onset of labour. This pattern was not exhibited in any of the other sensitivity analyses (see Appendix I for Forest plots for sensitivity analyses).

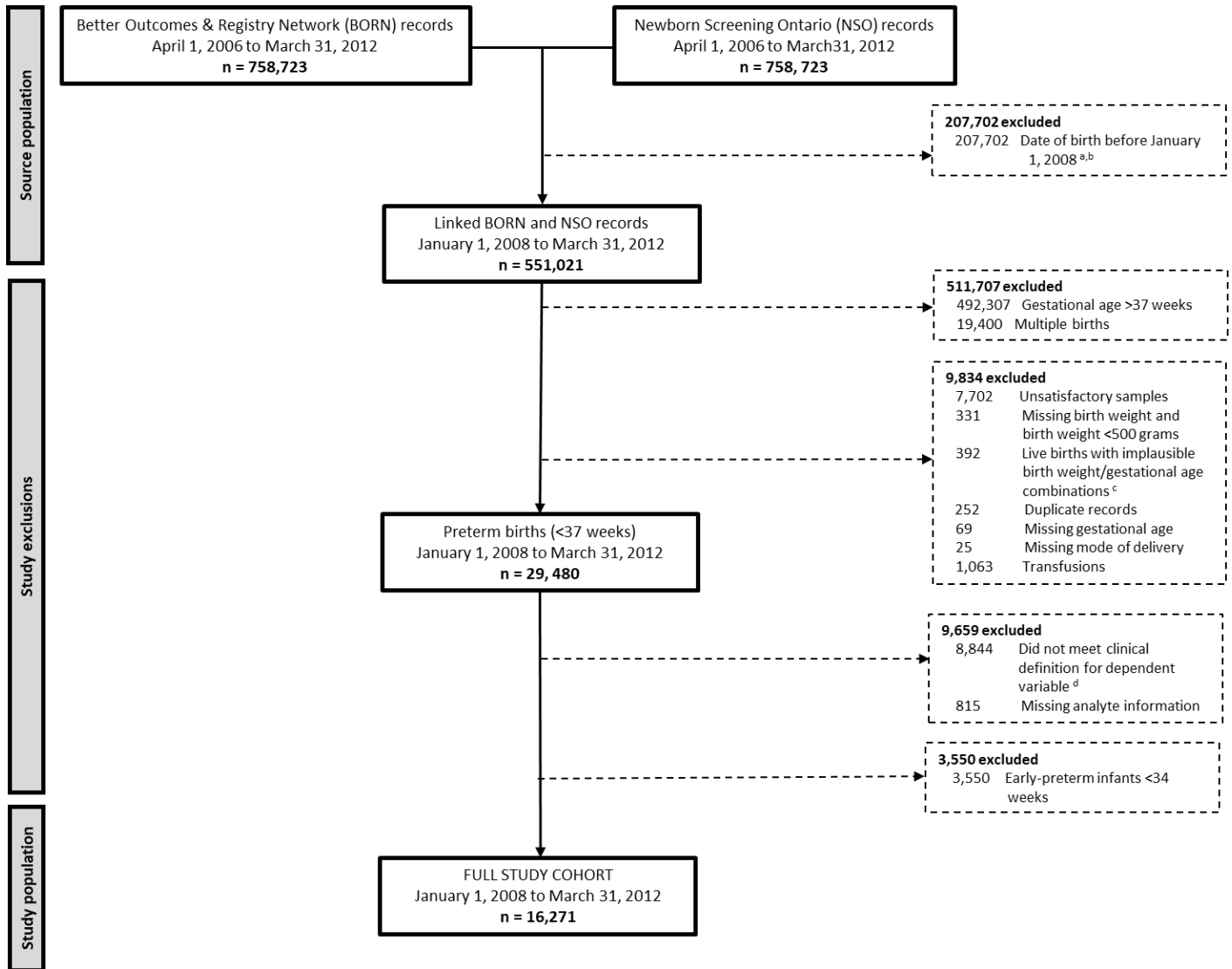
4.2 Objective 2

4.2.1 Characteristics of the Study Population

We included a total of 16,271 preterm infants in our analyses, following study exclusions (Figure 4.10). Of these, 4,405 infants (27.1%) were born at preterm gestation following PPRM, and 11,866 infants (72.9%) were born at preterm gestation following spontaneous onset of labour. The distribution of infant sex, TPN, and mode of delivery was similar across preterm birth subtypes. However, the proportions within each gestational age groups differed. Among infants born following PPRM, 44.9% were born at 36 weeks' gestation compared to 60.6% of infants born following spontaneous onset of labour. Preterm infants born following PPRM had a blood spot collection later compared with those

born following spontaneous onset of labour (62.1 hours, IQR: 31.8-75.4, and 56.0 hours, IQR: 27.1-63.0, respectively; Table 4.4).

Figure 4.10 Study flow diagram for Objective 2



^a Prior to 2008, analytes including biotinidase, 17-OHP, and fetal-to-adult hemoglobin level were not collected. However, from 2008 onwards biotinidase and 17-OHP were collected. Therefore, to ensure a large enough sample size and the maximum number of available analytes collected, we restricted the study period to 2008 and excluded fetal-to-adult hemoglobin level from analysis.

^b The exclusions of births prior to 2008 was applied to derive the source population. Additional objective-specific exclusions were applied afterwards (as demonstrated by the headings on the left).

^c Algorithm similar to that described by Basso and Wilcox⁹⁰ to clean gestational age values on live birth records was used.

^d Was not considered spontaneous onset of labour or preterm premature rupture of membranes.

Table 4.4 Baseline characteristics of singleton preterm live births in Ontario between 2008 and 2012 by preterm birth subtype

Characteristics	Preterm birth subtype (n=16, 271)			
	Preterm premature rupture of membranes (PPROM)		Spontaneous onset of labour prior to membrane rupture	
	n	%	n	%
All infants	4,405	100	11,866	100
Infant sex				
Female	1,952	44.3	5,188	43.7
Male	2,453	55.7	6,680	56.3
Gestational age (completed weeks)				
Mean \pm SD	35.2 \pm 0.8		35.5 \pm 0.7	
34	897	20.3	1,570	13.3
35	1,533	34.8	3,100	26.1
36	1,975	44.9	7,196	60.6
Birth weight (grams)				
Mean \pm SD	2,652.7 \pm 396.9		2,825.0 \pm 416.7	
<2,500	1,560	35.4	3,534	29.8
2,500-2,999	2,015	45.8	5,394	45.4
3,000-3,499	742	16.8	2,473	20.8
\geq 3,500	88	2.0	465	4.0
Age at blood spot collection (hours)				
Median (IQR)	62.1 (31.8 – 75.4)		56.0 (27.2 – 63.0)	
\leq 72	3,227	73.2	9,410	79.3
>72	1,178	26.8	2,456	20.7
Missing	0	0.0	0	0.0
Any total parenteral nutrition ^a				
No	4,348	98.6	11,758	99.1
Yes	57	1.4	108	0.9
Mode of delivery				
Vaginal	3,413	77.4	9,290	78.3
Cesarean	992	22.6	2,576	21.7

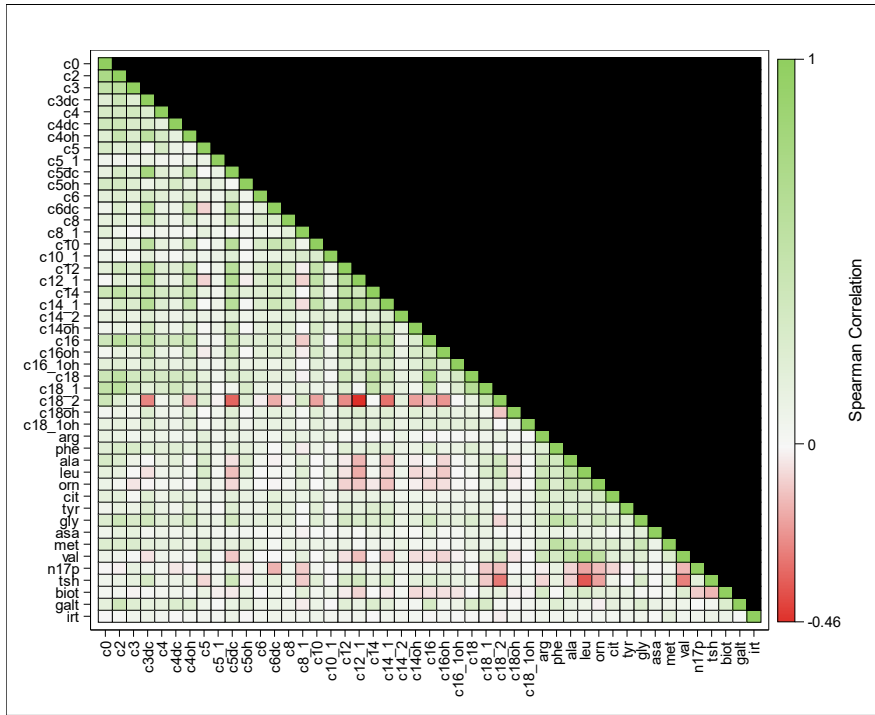
IQR: inter-quartile range; SD: standard deviation

^a Total parenteral nutrition alone or in combination with other infant feeding method.

4.2.2 Heat Map Analysis

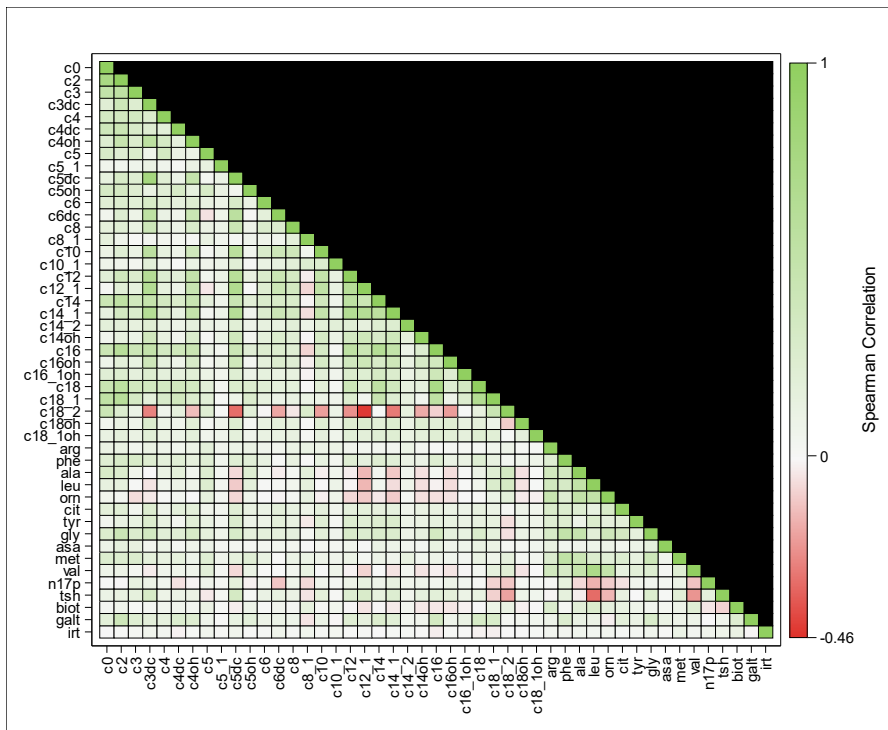
Among infants born preterm following PPRM, the strongest positive correlation was between free carnitine (C-0) and acetylcarnitine (C-2) ($\rho = 0.743$), while the strongest negative correlation was between octadecadienoylcarnitine (C-18:2) and dodecenoylcarnitine (C-12:1) ($\rho = -0.456$) (See Figure 4.11). Among those born following spontaneous onset of labour, the strongest positive correlation was also between free carnitine (C-0) and acetylcarnitine (C-2) ($\rho = 0.762$), while the strongest negative correlation was between octadecadienoylcarnitine (C-18:2) and dodecenoylcarnitine (C-12:1) ($\rho = -0.429$) (see Figure 4.12).

Figure 4.11 Correlation heatmaps for preterm infants born following PPROM



Darker green areas indicate higher levels of positive correlation between analytes, and darker red areas indicate higher levels of negative correlation between analytes.

Figure 4.12 Correlation heatmaps for preterm infants delivered following spontaneous onset of labour



Darker green areas indicate higher levels of positive correlation between analytes, and darker red areas indicate higher levels of negative correlation between analytes.

4.2.3 Partial Correlation analysis

The five strongest partial Spearman correlations between analytes or analyte ratios and PPROM were glutarylcarnitine (C-5DC) : octadecadienoylcarnitine (C-18:2), acetylcarnitine (C-2) : octadecadienoylcarnitine (C-18:2), dodecenoylcarnitine (C-12:1) : valine, methylmalonylcarnitine (C-4DC): leucine, and glycine : valine. The full results of the partial Spearman correlation analysis are presented in Appendix J.

4.2.4 Overall Model Performance

Discrimination

Table 4.5 summarizes the modelling performance metrics. With or without cross-validation, the discrimination for Model 1, which comprised the clinical variables only, performed the same with a c-statistic of 0.586, indicating a moderate fit. The discrimination of Model 2, which included 47 individual analytes and the top 100 ranked analyte ratios decreased slightly after cross-validation with a c-statistic of 0.578. With the inclusion of the clinical variables, the analytes and the analyte ratios, Model 3 performed the best with a cross-validated c-statistic of 0.605, though still in the moderate range.

Table 4.5 C-statistics for discrimination performance for Objective 2

Model		C-statistic (without cross- validation)	C-statistic (with cross- validation)
Model 1 ^a	Age at collection, sex, gestational age, birth weight, mode of delivery ^d	0.586	0.586
Model 2 ^b	47 individual analytes and 100 top ranked partial Spearman correlations for analyte ratios ^d	0.560	0.578
Model 3 ^c	Total parenteral nutrition, age at collection, sex, gestational age, birth weight, 47 individual analytes and 100 top ranked partial Spearman correlations for analyte ratios ^d	0.589	0.605

^a Model including the clinical variables only

^b Model including the analytes and analyte ratios only

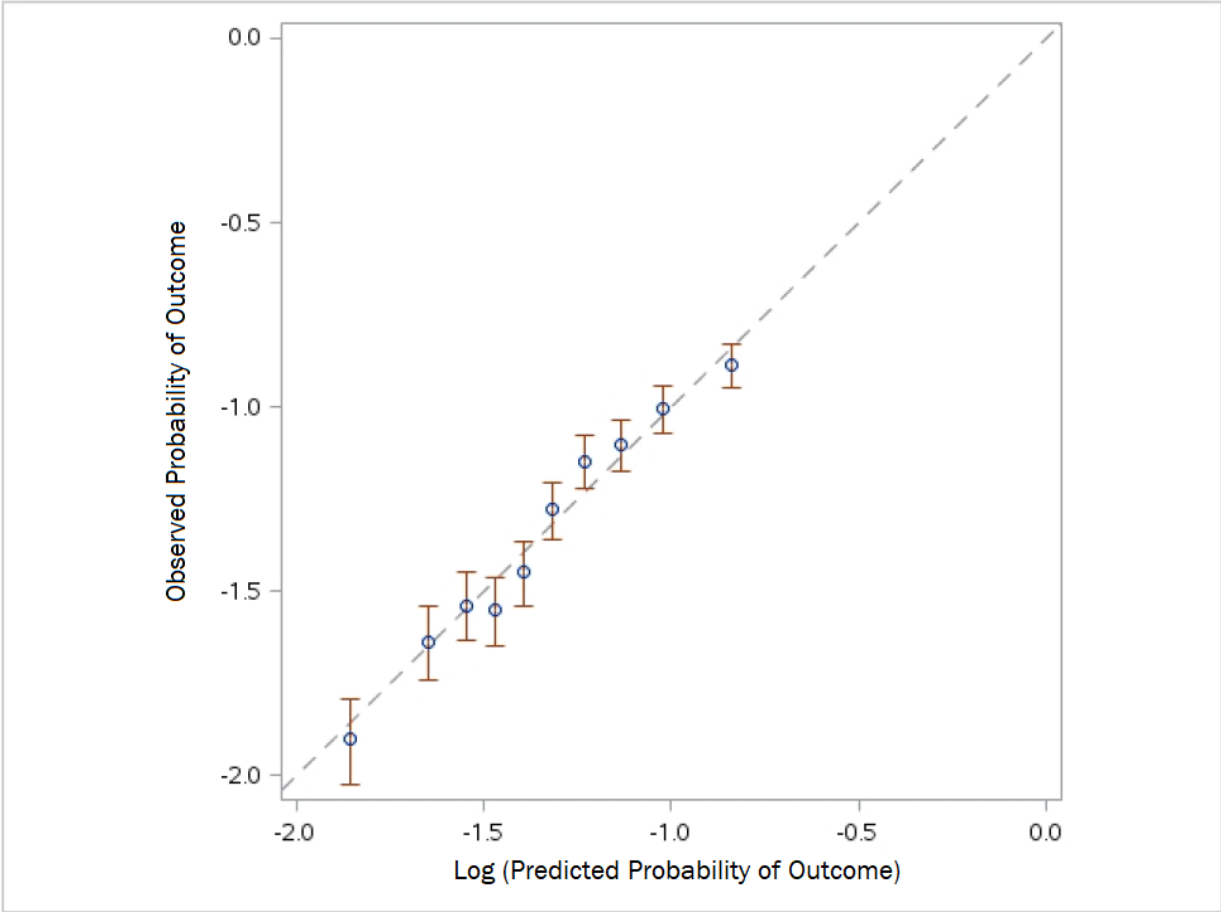
^c Model with a combination of the clinical variables and analytes and analyte ratios

^d These variables were covariates that were entered into the model, not necessarily the covariates that were retained following automated backward selection

Calibration

The calibration plot suggested that the predicted probability for PPRM was very similar to the observed probability of PPRM. The reference line captures the 95% CIs for the observed probability, indicating that the model (Model 3) was well-calibrated (Figure 4.13).

Figure 4.13 Decile calibration plots for preterm infants



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Association of newborn screening analytes with preterm birth subtype

The screening analytes that had the strongest positive associations with delivery following PPROM were acyl-carnitines and amino acids and ratios of these analytes, specifically valine and leucine.

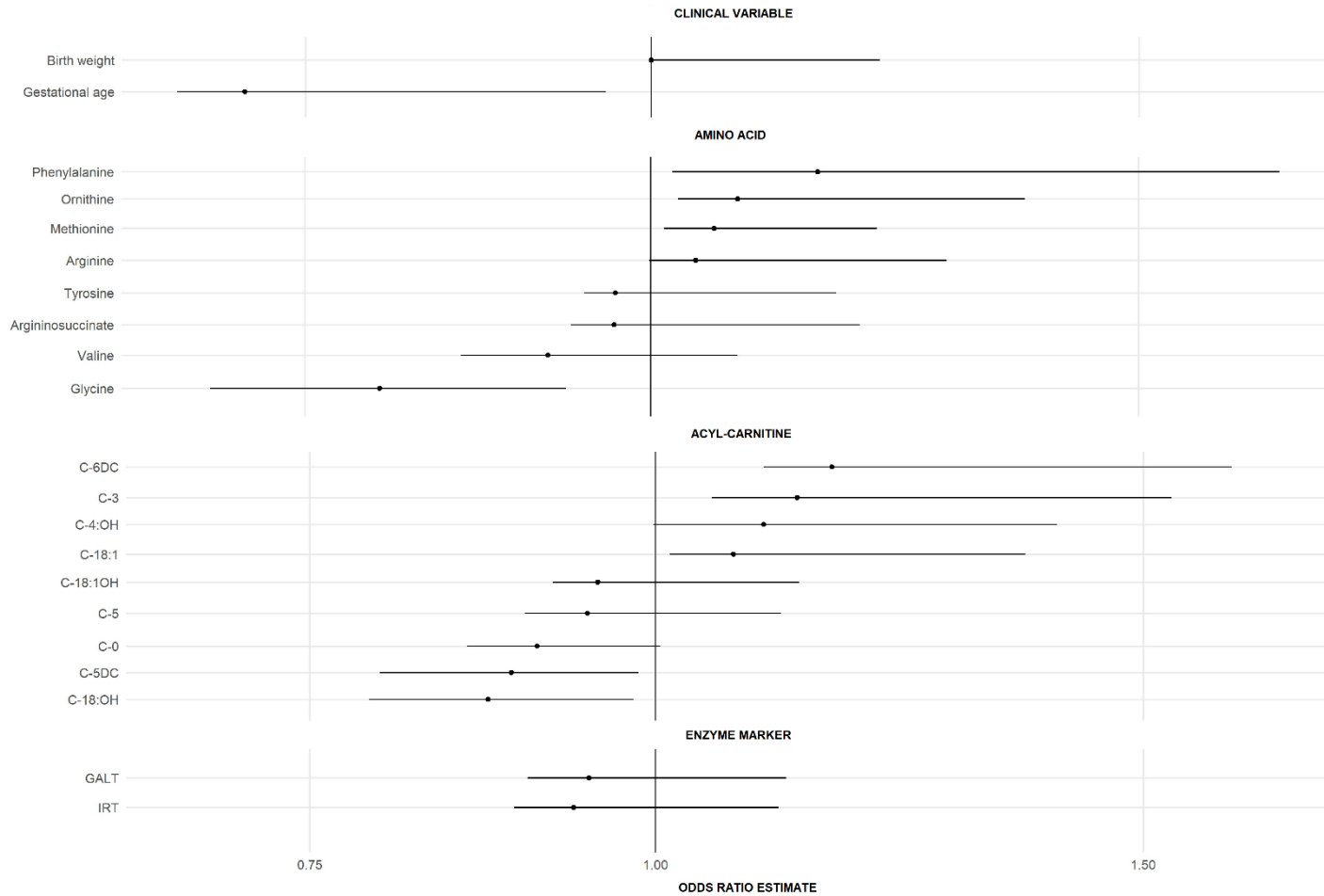
As compared with preterm infants born following spontaneous onset of labour, a 1 standard deviation increase in each of the following analytes was associated with an increase in the odds of preterm delivery following PPROM: C-5DC : C-18:2 (OR: 1.46; 95% CI: 1.20-4.83), C-2 : C-18:2 (OR: 1.22; 95% CI: 1.06-1.95), C-12:1 : Valine (OR: 1.16; 95% CI:1.02-1.64), C-6DC (OR: 1.16; 95% CI: 1.09-1.62), phenylalanine (OR: 1.15; 95% CI: 1.02-1.69), C-3DC : biotinidase (1.13; 95% CI: 1.03-1.41), C-3 (OR: 1.13; 95% CI: 1.05-1.54), C-18:OH : Leucine (OR: 1.13; 95% CI: 1.02-1.49), C-12:1 : Valine (OR: 1.10; 95% CI: 1.00-1.53), and TSH : biotinidase (OR: 1.10; 95% CI: 1.02-1.55) (see Figure 4.14a and 4.14b). These analytes were the top 10 analytes that were positively associated with preterm delivery following PPROM.

Newborn screening analytes that were most strongly inversely associated with preterm birth following PPROM were ratios of acyl-carnitines and amino acids, specifically glycine. Ratios of biotinidase with acyl-carnitines were also inversely associated with PPROM. As compared with delivery following spontaneous onset of labour , a 1 standard deviation increase in the following analytes was associated with a decrease in the odds of preterm birth following PPROM: glycine (OR: 0.80; 95% CI: 0.69-0.93), C-3DC : C-18:2 (OR: 0.83; 95% CI: 0.67-0.93), C-12:1 : biotinidase (OR: 0.88; 95% CI: 0.78-1.00), phenylalanine : glycine (OR: 0.88; 95% CI: 0.78-0.98), and C-4DC : biotinidase (OR: 0.89; 95% CI: 0.83-0.99) (see Figure 4.14a and 4.14b). These analytes were the top 10 analytes that were inversely associated with PPROM.

Only one clinical variable was significantly associated with preterm delivery following PPROM.

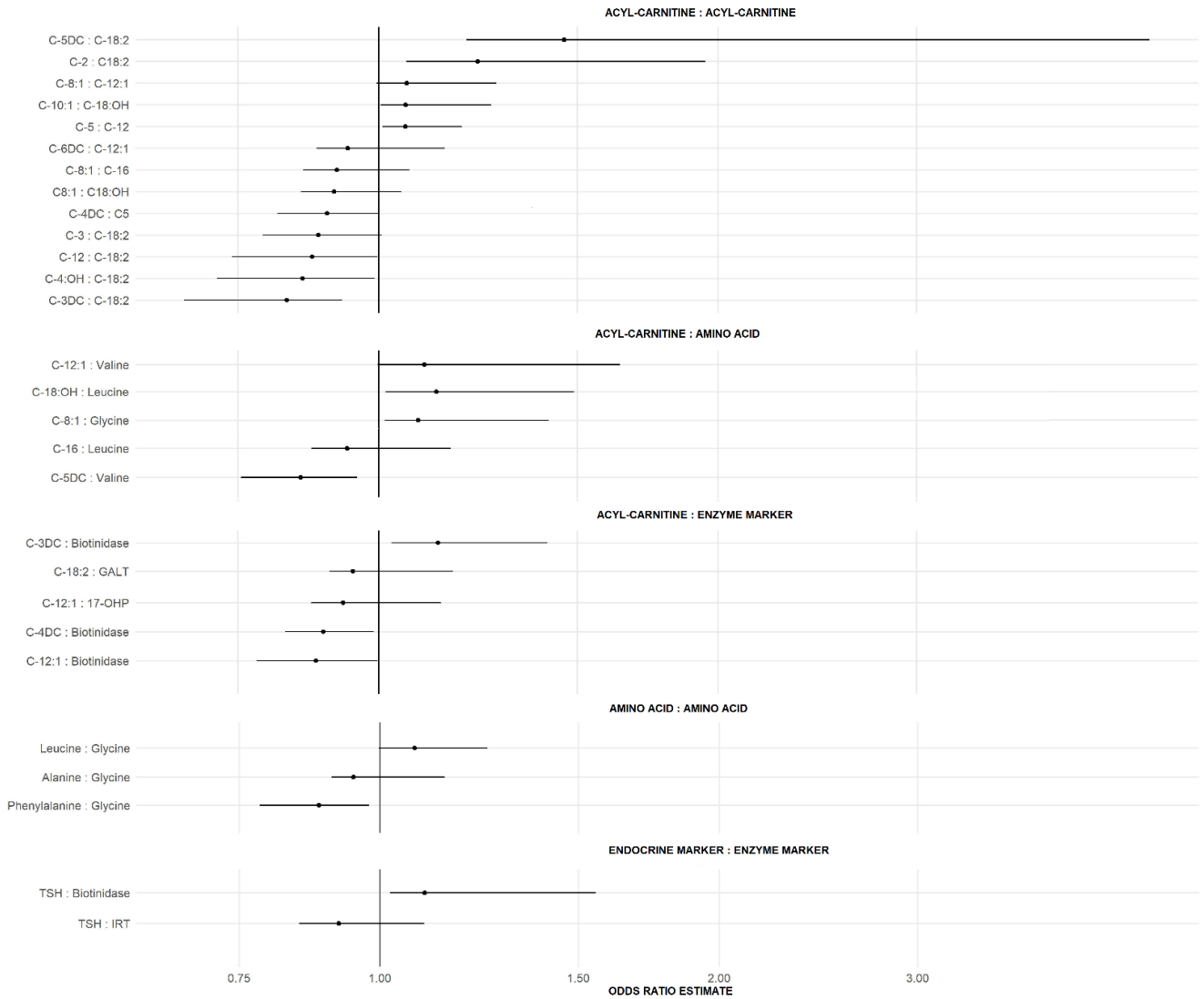
Gestational age was significantly inversely associated with PPROM, with an odds of 0.71 (95% CI: 0.67-0.96).

Figure 4.14a Individual analytes and clinical variables in the final model for preterm infants



Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Figure 4.14b Analyte ratios in the final model for preterm infants



Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

4.2.5 Overall Model Performance from Sensitivity Analyses

Discrimination

In a sensitivity analysis that included infants who underwent a transfusion and/or received TPN, the cross-validated c-statistic did not change appreciably. When we restricted the study period to 2010 and included the fetal-to-adult Hb ratio, the results for the cross-validated c-statistic decreased very slightly to 0.600. When we included the remaining analyte ratios until the maximum number of parameters was reached, the cross-validated c-statistic was 0.614 (Table 4.6).

Table 4.6 C-statistics for sensitivity analyses for Objective 2

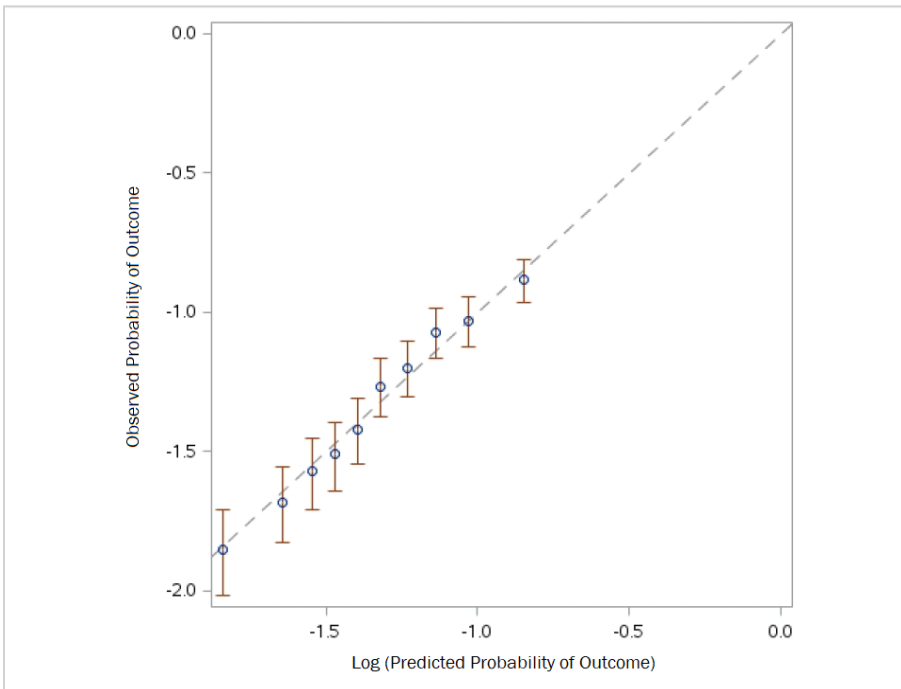
Model	C-statistic (without cross- validation) ^a	C-statistic (with cross- validation) ^a
Original analyses	0.589	0.605
Preterm infants including transfusions	0.589	0.605
Preterm infants born between 2010 and 2012	0.590	0.600
Preterm infants, including TPN in model	0.589	0.605
Inclusion of remaining ranked analyte ratios based on Spearman's correlation until the maximum number of parameters was reached	0.577	0.614

^a The c-statistic shown was obtained from the model with the combination of clinical variables and individual analytes and analyte ratios

Calibration

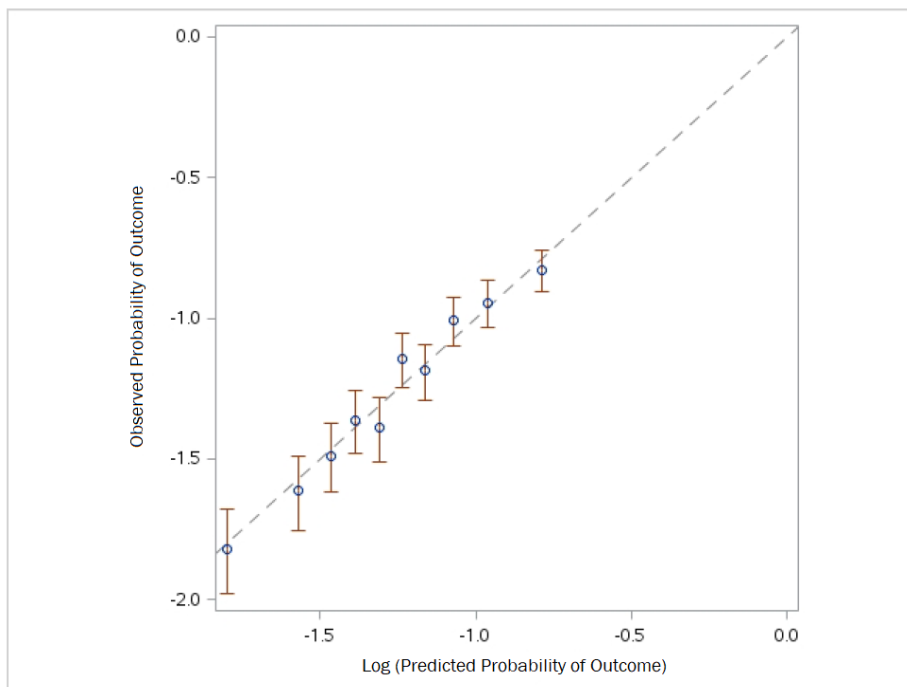
The calibration plots across three of the four sets of sensitivity analyses exhibited high calibration as the 95% confidence interval included the reference line data, indicating that the observed probability of PPRM closely matched the predicted probability of PPRM (see Figure 4.15 – Figure 4.17). The calibration in the fourth sensitivity analyses improved with increasing decile of predicted probability, agreeing closely with observed probability (Figure 4.18).

Figure 4.15 Decile calibration plots for preterm infants, including those who underwent transfusions



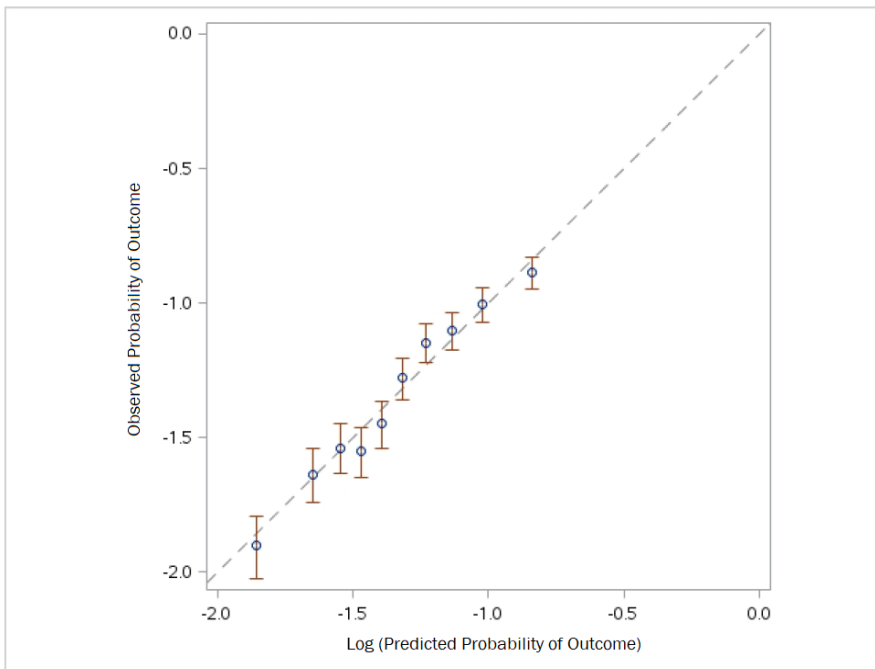
The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.16 Decile calibration plots for preterm infants born between 2010 and 2012



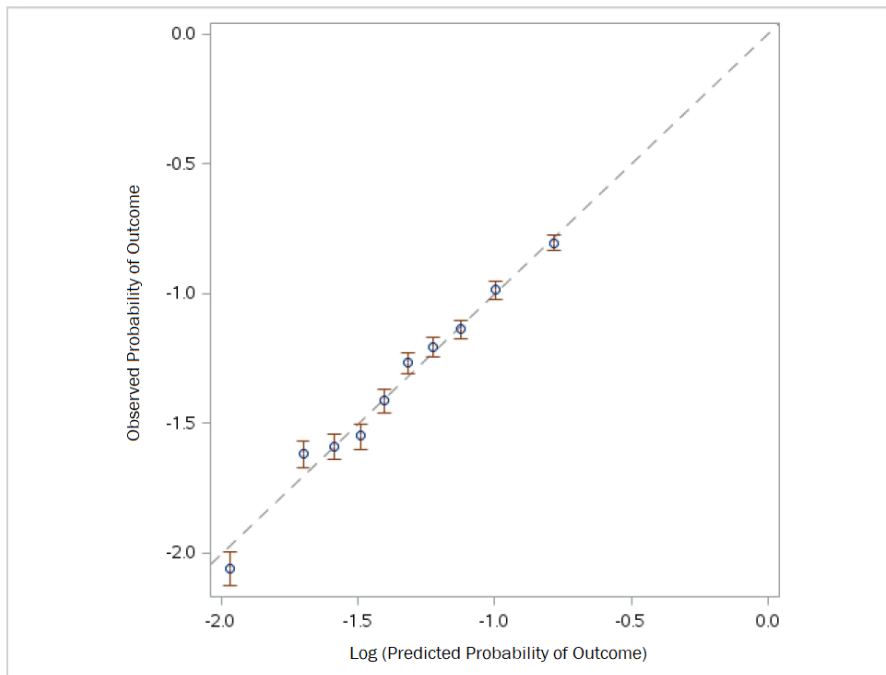
The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.17 Decile calibration plots for preterm infants, with the inclusion of TPN in the model



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Figure 4.18 Decile calibration plots after inclusion of remaining ranked analyte ratios until the maximum number of parameters was reached



The predicted probability of outcome has been log transformed. The calibration plot graphs the observed probability of outcome by the log (predicted probability of outcome). The 45° line represents the reference line. Data points that lie on the reference line indicate perfect calibration.

Association of newborn screening analytes with preterm birth subtype from sensitivity analyses

As seen in the main analysis findings, a combination of mostly acyl-carnitines and amino acids as well as enzyme markers, notably biotinidase was observed to be positively associated with PPROM across all sensitivity analyses. Notable differences were observed in the sensitivity analysis when we restricted the study period to 2010 onwards in order to include the fetal-to-adult Hb ratio, which only became available in NSO in 2010. We observed that a 1 standard deviation increase in fetal-to-adult Hb ratio, as well as the ratio of Hb to C-12 was positively associated with PPROM— OR: 2.35 (95% CI: 1.06-5.18) and 1.42 (95% CI: 1.19-1.70), respectively. Conversely, a 1 standard deviation increase in the ratio of fetal-to-adult Hb to C-4DC was strongly inversely associated with PPROM with an odds ratio of 0.72 (95% CI: 0.60-0.87) (See Appendix L for forest plots for sensitivity analyses).

CHAPTER 5. DISCUSSION

5.1 Statement of Principal Findings

Objective 1: In this study, we aimed to examine whether there were any association between newborn screening analytes and mode of delivery in term infants born to low-risk women. Our findings demonstrated that the newborn screening profiles of these infants differed by mode of delivery. In particular, a model with a combination of amino acids, acyl-carnitines, enzyme and endocrine markers, as well as, clinical variables performed well in distinguishing infants born by planned cesarean delivery from infants born by vaginal delivery following spontaneous onset of labour, over a model with clinical variables alone. The c-statistic for the final models suggested good discrimination in the full study population of all term infants (≥ 37 weeks; c-statistic=0.812) and similar discrimination was observed in sensitivity analyses, except when restricting the analyses to full-term infants (≥ 40 weeks) where the c-statistic indicated moderate fit (c-statistic=0.605).

Objective 2: In this study, we examined whether newborn screening analytes profiles were associated with clinical subtypes of preterm birth (i.e., preterm birth following the spontaneous onset of labour or following PPRM), independent from the actual mode of birth (i.e., whether or not the infant was delivered vaginally or by cesarean). Our model including clinical variables only suggested moderate performance (c-statistic=0.586) in distinguishing the different preterm birth subtypes. Although a combination of newborn screening analytes as well as clinical variables improved the model fit to 0.605, the gain was modest and suggested only a small incremental value in using acyl-carnitines, amino acids, enzyme marker, and endocrine markers, over clinical variables alone.

5.2 Interpretation of Findings

Objective 1: Our findings suggest that newborn screening analytes are associated with mode of delivery in low-risk infants born at term gestation.^{55,56} For example, we found that several amino acids and acyl-carnitines were positively associated with planned cesarean delivery, while several combinations of amino acids, acyl-carnitines, enzyme and endocrine markers were more inversely associated with planned cesarean delivery.

In particular, thyroid stimulating hormone (TSH) and galactose-1-phosphate-uridylyltransferase (GALT) were the screening analytes with the strongest inverse association with planned cesarean delivery. Several studies determined that levels of TSH in cord blood were significantly increased in infants born by vaginal delivery following spontaneous onset of labour compared to infants born by planned cesarean delivery.^{55,103,104} During vaginal delivery, the fetus is exposed to the stress of uterine contraction resulting in elevated levels of catecholamines, specifically noradrenaline, which can regulate the secretion of TSH from the fetal pituitary gland resulting in a rapid increase in TSH after delivery.¹⁰³ Our results support these findings, although we investigated blood obtained from the infant by heel prick, as opposed to cord blood TSH. In contrast, two studies showed no difference in neonatal TSH levels according to mode of delivery.^{105,106} However, one study included obtained blood samples from the infant at approximately 2.4 days after birth, when the postnatal surge of TSH would most likely have subsided.¹⁰⁶ In the second study, the cesarean delivery group could have included those that had no labour and those that had cesarean delivery following labour as information on when labour started was not available. Therefore, the process of labour could have started within this group which could explain why there was no difference observed in TSH among those born by vaginal delivery and cesarean

delivery.¹⁰⁵ In addition, the study did not exclude women with obstetrical or medical complications, which could have led to confounding by indication in that study.

Additionally, we observed an inverse association between the amino acid, alanine, and planned cesarean delivery. A previous study investigated amino acid levels in relation to mode of delivery by collecting blood from the mother and umbilical cord blood found higher levels of alanine in infants delivered vaginally following spontaneous onset of labour compared to those born by a scheduled cesarean delivery.¹⁰⁷ During the vaginal delivery process, the uterine and skeletal muscle contractions are more likely to be activated during labour and this could help to explain the increase in concentration of alanine.¹⁰⁸ This suggestion is further supported by evidence showing that levels of alanine are also significantly elevated among athletes after endurance exercise due to proteolysis.¹⁰⁹

We also observed an inverse association between acyl-carnitines and planned cesarean delivery, especially longer chain acyl-carnitines such as C-14:1. Higher concentrations of acyl-carnitines have been observed in plasma and tissue in cases of stress, likely due to disruptions in metabolic pathways.¹¹⁰ For example, levels of longer chain acyl-carnitines and other acyl-carnitines rise in the blood and muscle during exercise, as the demand for mitochondrial energy is increased.¹¹⁰⁻¹¹² This could explain the higher levels of acyl-carnitines with vaginal delivery observed in our study as the process of labour and the stress of vaginal delivery on the neonate can elicit the same physiological responses as the stress of exercise on the body. This could explain the inverse association between acyl-carnitines and planned cesarean delivery observed in our study, although, we used heel prick blood samples obtained from the newborn and the previously described studies used plasma and myocytes. Even so, the distinct

pathways of fatty acid oxidation as well as metabolic and endocrine function described may help explain how newborn screening analytes are associated with mode of delivery.

Objective 2: Different analyte patterns associated with subtypes of preterm birth were mildly observed. For example, we found that several amino acids and fetal-to-adult hemoglobin (Hb) level were positively associated with delivery following PPROM. The etiology of PPROM is not well understood but has been associated with decreased collagen content of the membranes.⁷⁷ Reactive oxygen species (ROS) could be a source for the collagen damage leading to premature rupture of the membranes.¹¹⁵ In situations of oxidative stress, the formation of ROS is increased which may modify the strength and elasticity of collagen.¹¹⁶ Increased concentrations of markers of oxidative stress in tissue from deliveries with PPROM have been reported.^{117,118} As amino acids are markers of oxidative stress, this is one possible explanation for the positive association of amino acids with PPROM in our study.

In one sensitivity analysis in which we restricted our study population to infants born between 2010 and 2012 and additionally included the Hb ratio, we observed that the Hb ratio was an important analyte in the model. In particular, the analyte alone and its ratio with C-12 was positively associated with increased odds of preterm birth following PPROM, compared with spontaneous preterm birth. Conversely, the Hb ratio to C-4DC was inversely associated with PPROM, suggesting that C-4DC could be modifying and changing the direction of the association between Hb and PPROM.

In other studies, elevated levels of Hb have been reported in infants born preterm following the onset of spontaneous preterm labour (with intact membranes at labour onset) as well as in those following

PPROM.^{119,120} However, these studies compared delivery following PPROM and spontaneous onset of labour with infants who presented with the same clinical presentation but also had intra-amniotic infection/inflammation. This could explain the association of Hb ratio in both preterm birth subtypes because hemoglobin and its catabolic products have been associated with discoloured amniotic fluid¹²¹ which is also associated with intra-amniotic infection/ inflammation, both of which have been linked to both subtypes of preterm birth.¹²⁰ Even small amounts of intra-amniotic bleeding can serve as a medium for bacterial growth¹²² and may increase the risk for intra-amniotic infection and subsequent preterm parturition.

Alternatively, findings from *Wilson et al.*²¹ demonstrated that levels of Hb ratio vary by gestational age. With advancing gestational age, hemoglobin production shifts from HbF to HbA.¹²³ Therefore, we would expect to see that Hb ratio, as defined by $Hb (F + F1)/Hb (A + F + F1)$, to be lowest within term infants and higher within preterm infants.²¹

5.3 Strengths

Objective 1: A strength of this study was the large sample size. Even after study exclusions, we included approximately 90,000 infants born at term gestation (≥ 37 weeks) in the analysis, which provided substantial statistical power to perform the complex analyses needed to develop our models. In addition, prioritization of covariates and internal validation were employed to maximize the quality of statistical models developed. Due to the large size of the cohort, we were able to partition our dataset into separate training and validation datasets. By internally validating our model, we were able to protect against over-fitting our model.

Previous studies have compared vaginal delivery and cesarean delivery; however, the distinction between emergency cesarean section and planned cesarean section is not usually made and could influence any associations seen when comparing the two types of delivery.^{55,105,124} The BORN-Niday Perinatal database that we utilized provided rich clinical information from which we were able to reliably categorize planned cesarean delivery (indicated for breech presentation, repeat cesarean section at term and maternal request with no medical indication) prior to the onset of labour, as well as vaginal delivery following spontaneous onset of labour so that we could examine the association between analytes and solely the mode by which the infant was delivered. A data quality assurance study completed on the BORN-Niday Perinatal database determined that all the variables required to characterize mode of delivery had between 90 and 99% agreement when compared to patient chart records, indicating high validity and completeness in measurement and reporting of these variables.⁸² Any outcome misclassification is thus expected to be small, non-differential, and have had minimal impact on our study findings.

Objective 2: This study also had a relatively large sample size. We had statistical power to split the subtypes of preterm birth into those following spontaneous onset of labour and those following PPROM. Even after excluding all medically-indicated preterm births, we still had a sample size of approximately 16,000 preterm infants. This meant that we had the statistical power to employ complex techniques to prioritize covariates into our model prior to modelling. Additionally, we used internal validation techniques (the leave-one-out cross validation) to cross-validate our model, to ensure the results were less susceptible to overfitting.

The data sources that we used in our study provided valuable information for us to sufficiently characterize clinical subtypes of preterm birth into those following spontaneous onset of labour and those following preterm premature rupture of membranes. While variability in obstetricians and research personnel in coding can occur, the clinical database from which we derived the subtypes of preterm birth has been found to be reliable and complete for core perinatal data variables such as gestational age and mode of delivery.⁸²

5.4 Limitations

As both objectives used the same data sources, they share many of the same limitations. We therefore begin with shared limitations across the two objectives, and then present objective-specific limitations below.

The data sources used in this study are generally not collected with the intent of performing research. As a result, missing data may present challenges. Infants who died before newborn screening were not included due to the unavailability of information on these events in the BORN-Niday database and NSO database. This can potentially introduce selection bias if the newborn screening analytes are associated with these early adverse neonatal outcomes. However, we were unable to evaluate this in my study. In addition, infants who died before newborn screening could be collected are more likely to have had compromised metabolic functions, which may have influenced the quantitative values of the analytes in ways that influenced our assessment.

Measurement of gestational age in the database only captures the number of completed weeks of gestation. The source of the gestational age estimate, whether it is based on last menstrual period or ultrasound, is not documented in the database; however, early ultrasound assessment is the dating method that was used for most pregnancies during our study period in Ontario.¹²⁵ We used an algorithm to identify records for implausible birth weight and gestational age combinations as part of data cleaning.⁹⁰ While gestational age in weeks was collected in the data sources we used and included in our models, residual confounding could still exist. Gestational age is highly correlated with newborn screening analyte levels.¹² It is also highly correlated with type of delivery. For instance, in Objective 1 approximately half of infants born by planned cesarean delivery were born at 38 weeks' gestation vs. 17.9% of infants born following vaginal delivery. In Objective 2, among infants born following PPROM, 44.9% were born at 36 weeks of completed gestation whereas among infants born following spontaneous onset of labour, 60.6% were born at 36 weeks' gestation. Because of this, we forced gestational age into our models for both objectives. In sensitivity analyses for Objective 1, we tried to account for GA even more closely by additionally stratifying our analyses to early-term infants (37-39 weeks) and full-term infants (≥ 40 weeks). Similarly, for Objective 2, we stratified gestational age into early-preterm infants (< 34 weeks) and late-preterm infants (34-36 weeks) to more closely account for GA. However, we ultimately analyzed only the late-preterm group of infants (34-36 weeks) because the sample size for the early-preterm group of infants precluded the type of complex statistical analyses we were conducting. Most preterm births occur between 34 and 36 weeks, providing a greater sample size.¹⁸ However, despite these strategies, it is still possible that residual confounding could exist even within the gestational age strata that we created. Although the impact would be small, we suspected that this would have strengthened the associations between newborn screening analytes and mode of delivery or preterm birth subtypes, either positively or inversely because previous research has found that newborn screening analyte levels were highly correlated with gestational age.¹²

Our study population excluded infants who screened positive, including those who were false-positives. Ideally, we would have included the latter infants in the study population, as they were not ultimately diagnosed with a metabolic disorder. However, we were not able to include them because this exclusion had already been applied to the dataset that comprised the source population for our study. Although the infants with false-positive screening results clearly had some analyte levels outside of normal ranges, which could have been informative in our study, we anticipate that this would have been a small number relative to the very large study population and would not have introduced any strong bias into our results.

Another challenge of our study was that interpretation of our results was limited and difficult. Firstly, dried blood samples were collected after the delivery has occurred, making it difficult to determine the directionality of the association, as the event has happened retrospectively. Additionally, while we accounted for approximate interaction terms newborn screening analytes, the calculation of analyte ratios has rendered the interpretation of our results difficult as the analytes are not mutually dependent and must be interpreted in relation to one another. The results from our study should thus be seen as hypothesis-generating.

Objective 1: We tried to select a population that would ensure that women who had a vaginal or a cesarean delivery were as comparable as possible so that solely the mode of delivery could be compared, not any clinical indication for mode of delivery. To do so, we excluded women with a preterm birth, pre-existing health problems, obstetric complications, intrapartum complications, and among those with a cesarean delivery, only those with one of three specific clinical indications for cesarean delivery so that we would have low-risk deliveries. However, it is possible that some level of disease

states could still exist within our study population due to under-diagnosis or lack of documentation in the medical record. We believe that this would be limited to a small subset of the women in our study and would not have a large influence on the newborn screening analytes.

Objective 2: Our understanding of preterm birth has been complicated by the difficulty in defining and classifying preterm birth.^{17,75} Existing phenotypic classifications include subdivisions by gestational age, clinical presentation, and presumed pathophysiological pathways.¹⁷ Additionally, preterm birth is known to have a multitude of heterogeneous risk factors that can involve medical, genetic, social, and environmental components⁵, adding a layer of complexity in the classification of preterm birth subtypes. Given that the etiological pathway of subtypes of preterm birth is multifactorial, it could vary by individual interpretation of obstetric providers and research personnel. A recent study found incorrect classification of preterm birth subtypes 5–15% of the time, even with experienced research personnel.¹²⁶ This could have resulted in non-differential misclassification bias as we could have incorrectly classified infants into either spontaneous onset of labour or PPRM due to misinterpretation of coding. However, this would mostly likely have diluted any associations we observed.

5.5 Conclusions

Due to our reliance on newborn screening to identify vulnerable children in the early days of life to prevent severe health problems, a thorough investigation into the influences of environmental and biological factors, including mode of delivery and subtypes of preterm birth, is warranted for understanding variability in newborn screening analytes. Our interpretation of newborn screening results, particularly when being used beyond their traditional application, such as in postnatal

gestational age estimation²¹ and predictive models for other neonatal morbidities²², can also be improved and informed by the results from this thesis. This thesis project aimed to determine whether there was an association between newborn screening analytes profiles and mode of delivery among infants born at term gestation to low-risk women (Objective 1) and birth subtypes among infants born at preterm gestation (Objective 2).

Objective 1: The results of our study suggest that the metabolic profiles of infants born by planned cesarean delivery differ from those born by vaginal delivery following spontaneous onset of labour, demonstrating that newborn screening analytes had discriminatory ability to categorize modes of delivery. This suggests that modes of delivery should be carefully considered as potential confounders, effect modifiers or as stratification factors in which separate models might need to be developed according to mode of delivery, when planning modelling studies involving newborn screening metabolomics and prediction of future outcomes.

Objective 2: Results were less compelling with respect to differences in metabolic profiles among preterm infants born following spontaneous onset of labour and those born following PPRM. Specific to this objective, newborn screening analytes did not provide enough information to effectively categorize preterm birth subtypes. Thus, their inclusion as a covariate in future modelling studies does not appear to be as important.

REFERENCES

1. Hood L. Systems biology: Integrating technology, biology, and computation. *Mechanisms of Ageing and Development*. 2003. p. 9–16.
2. Gravett MG, Rubens CE, Nunes TM, GAPPs Review Group. Global report on preterm birth and stillbirth (2 of 7): Discovery Science. *BMC Pregnancy Childbirth*. 2010 Feb 23;10 Suppl 1:S2.
3. Romero R, Espinoza J, Gotsch F, Kusanovic JP, Friel LA, Erez O, et al. The use of high-dimensional biology (genomics, transcriptomics, proteomics, and metabolomics) to understand the preterm parturition syndrome. *BJOG An Int J Obstet Gynaecol*. 2006;113(Suppl 3):118–35.
4. Georgiou HM, Di Quinzio MKW, Permezel M, Brennecke SP. Predicting Preterm Labour: Current Status and Future Prospects. *Dis Markers*. 2015;2015:1–9.
5. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. Vol. 371, *The Lancet*. 2008. p. 75–84.
6. Newborn Screening Ontario. About screening. 2016. Available from: <https://www.newbornscreening.on.ca/en/about-screening/what-screening>
7. Ontario Ministry of Health and Long-Term Care. Newborn Screening. 2013. Available from: <http://www.health.gov.on.ca/en/pro/programs/newborn/>
8. Wilson K, Kennedy SJ, Potter BK, Geraghty MT, Chakraborty P. Developing a National Newborn Screening Strategy for Canada. *Health Law Rev*. 2010;18(2).
9. Schreiner F, Brack C, Salzgeber K, Vorhoff W, Woelfle J, Gohlke B. False negative 17-hydroxyprogesterone screening in children with classical congenital adrenal hyperplasia. *Eur J Pediatr*. 2008;167(4):479–81.
10. Ryckman KK, Berberich SL, Shchelochkov OA, Cook DE, Murray JC. Clinical and environmental influences on metabolic biomarkers collected for newborn screening. *Clin Biochem*. 2013;46(1–2):133–8.
11. Jelliffe-Pawlowski LL, Norton ME, Baer RJ, Santos N, Rutherford GW. Gestational dating by metabolic profile at birth: A California cohort study. *American Journal of Obstetrics and Gynecology*. 2016. p. 511.e1-511.e13.
12. Wilson K, Hawken S, Potter BK, Chakraborty P, Walker M, Ducharme R, et al. Accurate prediction of gestational age using newborn screening analyte data. *Am J Obstet Gynecol*. 2016;214(4):513.e1-513.e9.
13. Neu J, Rushing J. Cesarean Versus Vaginal Delivery: Long-term Infant Outcomes and the Hygiene Hypothesis. *Clinics in Perinatology*. 2011.
14. Syed M, Javed H, Yakoob MY, Bhutta ZA. Effect of screening and management of diabetes during pregnancy on stillbirths. *BMC Public Health*. 2011 Apr 13;11(Suppl 3):S2.
15. World Health Organization. Preterm birth. 2016. Available from: <http://www.who.int/mediacentre/factsheets/fs363/en/>

16. Blencowe H, Cousens S, Oestergaard MZ, Chou D, Moller A-B, Narwal R, et al. 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*. 2012;379(9832):2162–72.
17. Kramer MS, Papageorghiou A, Culhane J, Bhutta Z, Goldenberg RL, Gravett M, et al. Challenges in defining and classifying the preterm birth syndrome. Vol. 206, *American Journal of Obstetrics and Gynecology*. 2012. p. 108–12.
18. Savitz DA, Dole N, Herring AH, Kaczor D, Murphy J, Siega-Riz AM, et al. Should spontaneous and medically indicated preterm births be separated for studying aetiology? *Paediatr Perinat Epidemiol*. 2005;19(2):97–105.
19. Ananth C V, Vintzileos AM. Epidemiology of preterm birth and its clinical subtypes. *J Matern Fetal Neonatal Med*. 2006;19(12):773–82.
20. Muglia LJ, Katz M. The Enigma of Spontaneous Preterm Birth. *N Engl J Med*. 2010 Feb 11;362(6):529–35.
21. Wilson K, Hawken S, Murphy MSQ, Atkinson KM, Potter BK, Sprague A, et al. Postnatal Prediction of Gestational Age Using Newborn Fetal Hemoglobin Levels. *EBioMedicine*. 2017 Feb 1;15:203–9.
22. Fell DB, Hawken S, Wong CA, Wilson LA, Murphy MSQ, Chakraborty P, et al. Using newborn screening analytes to identify cases of neonatal sepsis. *Sci Rep*. 2017 Dec 21;7(1):18020.
23. Evans GA. Designer science and the “omic” evolution. *Nat Biotechnol*. 2000 Feb;18(2):127.
24. Shankar R, Cullinane F, Brennecke SP, Moses EK. Applications of proteomic methodologies to human pregnancy research: A growing gestation approaching delivery? *Proteomics*. WILEY-VCH Verlag; 2004. p. 1909–17.
25. Fanos V, Atzori L, Makarenko K, Melis GB, Ferrazzi E. Metabolomics application in maternal-fetal medicine. Vol. 2013, *BioMed Research International*. 2013.
26. Auray-Blais C, Raiche E, Gagnon R, Berthiaume M, Pasquier JC. Metabolomics and preterm birth: What biomarkers in cervicovaginal secretions are predictive of high-risk pregnant women? *Int J Mass Spectrom*. 2011;307(1–3):33–8.
27. Ghartey J, Bastek JA, Brown AG, Anglim L, Elovitz MA. Women with preterm birth have a distinct cervicovaginal metabolome. *American Journal of Obstetrics and Gynecology*. 2015. p. 776.e1-776.e12.
28. Cecatti JG, Souza RT, Sulek K, Costa ML, Kenny LC, McCowan LM, et al. Use of metabolomics for the identification and validation of clinical biomarkers for preterm birth: Preterm SAMBA. *BMC Pregnancy Childbirth*. 2016;16(1):212.
29. Wilson K, Hawken S, Ducharme R, Potter BK, Little J, Thébaud B, et al. Metabolomics of prematurity: analysis of patterns of amino acids, enzymes, and endocrine markers by categories of gestational age. *Pediatr Res*. 2014;75(2):367–73. A
30. Imai M, Tani A, Saito M, Saito K, Amano K, Nisijima M. Significance of fetal fibronectin and cytokine measurement in the cervicovaginal secretions of women at term in predicting term

- labor and post-term pregnancy. *Eur J Obstet Gynecol Reprod Biol* . 2001 Jul;97(1):53–8.
31. Goldenberg RL, Goepfert AR, Ramsey PS. Biochemical Markers for the Prediction of Preterm Delivery . Vol. 34, *American Journal of Obstetrics and Gynecology*. 2005. p. 573–86.
 32. Goldenberg RL, Iams JD, Mercer BM, Meis PJ, Moawad A, Das A, et al. The Preterm Prediction Study: Toward a multiple-marker test for spontaneous preterm birth. *Am J Obstet Gynecol* . 2001 Sep;185(3):643–51.
 33. Shah J, Baxi B. Identification of biomarkers for prediction of preterm delivery. Vol. 30, *Journal of Medical Society*. 2016. p. 3–14.
 34. Heng YJ, Liong S, Permezel M, Rice GE, Di Quinzio MKW, Georgiou HM. Human cervicovaginal fluid biomarkers to predict term and preterm labor. *Front Physiol*. 2015;6(May):151.
 35. Knight A, Smith A. Epigenetic Biomarkers of Preterm Birth and Its Risk Factors. *Genes*. 2016;7(4):15.
 36. Menon R, Torloni MR, Voltolini C, Torricelli M, Merialdi M, Betran a P, et al. Biomarkers of spontaneous preterm birth: an overview of the literature in the last four decades. *Reprod Sci*. 2011 Oct 25;18(11):1046–70.
 37. Lucaroni F, Morciano L, Rizzo G, D' Antonio F, Buonomo E, Palombi L, et al. Biomarkers for predicting spontaneous preterm birth: an umbrella systematic review. *J Matern Neonatal Med* . 2017 Mar 8;1–9.
 38. Thorsen P, Schendel DE, Deshpande a D, Vogel I, Dudley DJ, Olsen J. Identification of biological/biochemical marker(s) for preterm delivery. *Paediatr Perinat Epidemiol*. 2001;15 Suppl 2:90–103.
 39. Yoon H-R. Screening newborns for metabolic disorders based on targeted metabolomics using tandem mass spectrometry. *Ann Pediatr Endocrinol Metab*. 2015 Sep;20(3):119–24.
 40. Therrell BL, Adams J. Newborn screening in North America. *Journal of Inherited Metabolic Disease*. 2007. p. 447–65.
 41. Therrell BL, Padilla CD, Loeber JG, Kneisser I, Saadallah A, Borrajo GJ, et al. Current status of newborn screening worldwide: 2015. *Semin Perinatol*. 2015;39(3):171–87.
 42. Padilla CD, Therrell BL. Newborn screening in the Asia Pacific region. *Journal of Inherited Metabolic Disease*. 2007. p. 490–506.
 43. Bodamer OA, Hoffmann GF, Lindner M. Expanded newborn screening in Europe 2007. *Journal of Inherited Metabolic Disease*. 2007. p. 439–44.
 44. Borrajo GJC. Newborn screening in Latin America at the beginning of the 21st century. *Journal of Inherited Metabolic Disease*. 2007. p. 466–81.
 45. Saadallah AA, Rashed MS. Newborn screening: Experiences in the Middle East and North Africa. *Journal of Inherited Metabolic Disease*. 2007. p. 482–9.
 46. Newborn Screening Ontario. Backgrounder on Newborn Screening Ontario . 2017. Available

from:https://www.newbornscreening.on.ca/sites/default/files/pdfs/nso_backgrounder_2017_final.pdf

47. Newborn Screening Ontario. Newborn Screening Ontario Data User Guide. 2016.
48. Singh RJ. Quantitation of 17-OH-progesterone (OHPG) for diagnosis of congenital adrenal hyperplasia (CAH). *Clin Appl Mass Spectrom*. 2010;603:271–7.
49. Newborn Screening Ontario. Newborn Screening Manual: A Guide for Newborn Care Providers. 2015;(January):1–60. Available from: <http://www.newbornscreening.on.ca/bins/index.asp>
50. Chace DH, Spitzer AR. Altered metabolism and newborn screening using tandem mass spectrometry: lessons learned from the bench to bedside. *Curr Pharm Biotechnol*. 2011 Jul;12(7):965–75.
51. Varness TS, Allen DB, Hoffman GL. Newborn screening for congenital adrenal hyperplasia has reduced sensitivity in girls. *J Pediatr*. 2005 Oct;147(4):493–8.
52. Pang S. Newborn screening for congenital adrenal hyperplasia. *Pediatr Ann*. 2003 Aug;32(8):516–23.
53. Allen DB, Hoffman GL, Fitzpatrick P, Laessig R, Maby S, Slyper A. Improved precision of newborn screening for congenital adrenal hyperplasia using weight-adjusted criteria for 17-hydroxyprogesterone levels. *J Pediatr*. 1997 Jan;130(1):128–33.
54. Merenstein DJ, Gatti ME, Mays DM. The association of mode of delivery and common childhood illnesses. *Clin Pediatr (Phila)*. 2011;50(11):1024–30.
55. Lagercrantz H. The good stress of being born. *Acta Paediatr Int J Paediatr*. 2016 Dec;105(12):1413–6.
56. Tribe RM, Taylor PD, Kelly NM, Rees D, Sandall J, Kennedy HP. Parturition and the perinatal period: Can mode of delivery impact on the future health of the neonate? *Journal of Physiology*. 2018 Apr 15.
57. Hillman NH, Kallapur SG, Jobe AH. Physiology of transition from intrauterine to extrauterine life. Vol. 39, *Clinics in Perinatology*. 2012. p. 769–83.
58. Vogl SE, Worda C, Egarter C, Bieglmayer C, Szekeres T, Huber J, et al. Mode of delivery is associated with maternal and fetal endocrine stress response. *BJOG An Int J Obstet Gynaecol*. 2006;113(4):441–5.
59. Reproductive Care Program of Nova Scotia. Best practices in the use of Cesarean sections in Nova Scotia . 2008. Available from: <http://rcp.nshealth.ca/publications/best-practices-use-cesarean-sections-nova-scotia>
60. Health Canada. Perinatal Health Indicators for Canada. Ottawa; 2000 . 87 p. Available from: http://publications.gc.ca/collections/collection_2018/aspc-phac/HP7-1-2017-eng.pdf
61. Mylonas I, Friese K. Indications for and Risks of Elective Cesarean Section. *Dtsch Ärzteblatt Int*. 2015 Jul 20;112(29–30):489–95.

62. Mozurkewich EL, Hutton EK. Elective repeat cesarean delivery versus trial of labor: A meta-analysis of the literature from 1989 to 1999. *American Journal of Obstetrics and Gynecology*. 2000. p. 1187–97.
63. Signore C, Klebanoff M. Neonatal Morbidity and Mortality After Elective Cesarean Delivery . Vol. 35, *Clinics in Perinatology. NIH Public Access*; 2008. p. 361–71.
64. Landon MB, Hauth JC, Leveno KJ, Spong CY, Leindecker S, Varner MW, et al. Maternal and Perinatal Outcomes Associated with a Trial of Labor after Prior Cesarean Delivery. *N Engl J Med*. 2004 Dec 16;351(25):2581–9.
65. Rossi AC, Dapos;Addario V. Maternal morbidity following a trial of labor after cesarean section vs elective repeat cesarean delivery: a systematic review with metaanalysis . Vol. 199, *American Journal of Obstetrics and Gynecology*. 2008. p. 224–31.
66. Hannah ME, Hannah WJ, Hewson SA, Hodnett ED, Saigal S, Willan AR. Planned caesarean section versus planned vaginal birth for breech presentation at term: A randomised multicentre trial. *Lancet*. 2000 Oct 21;356(9239):1375–83.
67. Hogle KL, Kilburn L, Hewson S, Gafni A, Wall R, Hannah ME. Impact of the international term breech trial on clinical practice and concerns: a survey of centre collaborators. *J Obstet Gynaecol Canada*. 2003;25(1):14–6.
68. Hofmeyr GJ, Hannah M, Lawrie TA. Planned caesarean section for term breech delivery. Vol. 2015, *Cochrane Database of Systematic Reviews*. 2015.
69. McCourt C, Weaver J, Statham H, Beake S, Gamble J, Creedy DK. Elective cesarean section and decision making: A critical review of the literature. Vol. 34, *Birth*. 2007. p. 65–79.
70. Lin HC, Xirasagar S. Maternal age and the likelihood of a maternal request for cesarean delivery: A 5-year population-based study. *American Journal of Obstetrics and Gynecology*. Mosby; 2005. p. 848–55.
71. American College of Obstetricians and Gynecologists. ACOG Committee Opinion No. 386: Cesarean delivery on maternal request . Vol. 110, *Obstetrics and Gynecology*. 2007. p. 1209–12.
72. Souza JP, Gülmezoglu AM, Lumbiganon P, Laopaiboon M, Carroli G, Fawole B, et al. Caesarean section without medical indications is associated with an increased risk of adverse short-term maternal outcomes: The 2004-2008 WHO Global Survey on Maternal and Perinatal Health. *BMC Med*. 2010 Dec 10;8(1):71.
73. Patterson DA, Winslow M, Matus CD. Spontaneous vaginal delivery . Vol. 78, *American Family Physician*. 2008.
74. Barros FC, Papageorghiou AT, Victora CG, Noble JA, Pang R, Iams J, et al. The Distribution of Clinical Phenotypes of Preterm Birth Syndrome. *JAMA Pediatr*. 2015;1–10.
75. Goldenberg RL, Gravett MG, Iams J, Papageorghiou AT, Waller SA, Kramer M, et al. The preterm birth syndrome: Issues to consider in creating a classification system. *Am J Obstet Gynecol*. 2012;206(2):113–8.

76. Di Renzo GC, Roura LC. Guidelines for the management of spontaneous preterm labor. Vol. 34, *Journal of Perinatal Medicine*. 2006. p. 359–66.
77. Medina TM, Hill DA. Preterm Premature Rupture of Membranes: Diagnosis and Management. *Am Fam Physician*. 2006;73(4):659–64.
78. Ananth C V., Vintzileos AM. Medically Indicated Preterm Birth: Recognizing the Importance of the Problem . Vol. 35, *Clinics in Perinatology*. 2008. p. 53–67.
79. Spong CY, Mercer BM, D’Alton M, Kilpatrick S, Blackwell S, Saade G. Timing of Indicated Late-Preterm and Early-Term Birth. *Obstet Gynecol*. 2011;118(2, Part 1):323–33.
80. American College of Obstetricians and Gynecologists. American College of Gynecology Committee Opinion. *American College of Obstetricians and Gynecologists*. Available from: <http://www.acog.org/Resources-And-Publications/Committee-Opinions/Committee-on-Obstetric-Practice/Medically-Indicated-Late-Preterm-and-Early-Term-Deliveries>
81. Newborn Screening Ontario. Newborn Screening Ontario Annual Report 2015. 2015.
82. Dunn S, Bottomley J, Ali A, Walker M. 2008 Niday perinatal database quality audit: Report of a quality assurance project. *Chronic Dis Can*. 2011 Dec;32(1):32–42.
83. Liu S, Liston RM, Joseph KS, Heaman M, Sauve R, Kramer MS. Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term. *CMAJ*. 2007 Feb 13;176(4):455–60.
84. Dunn S, Sprague AE, Grimshaw JM, Graham ID, Taljaard M, Fell D, et al. A mixed methods evaluation of the maternal-newborn dashboard in Ontario: dashboard attributes, contextual factors, and facilitators and barriers to use: a study protocol. *Implement Sci*. 2015.
85. Hannah ME, Hannah WJ, Hewson SA, Hodnett ED, Saigal S, Willan AR. Planned caesarean section versus planned vaginal birth for breech presentation at term: A randomised multicentre trial. *Lancet*. 2000 Oct 21;356(9239):1375–83.
86. Allen VM, O’Connell CM, Liston RM, Baskett TF. Maternal morbidity associated with cesarean delivery without labor compared with spontaneous onset of labor at term. *Obstet Gynecol* . 2003 Sep;102(3):477–82.
87. Kugelman A, Colin AA. Late Preterm Infants: Near Term But Still in a Critical Developmental Time Period. *Pediatrics*. 2013 Oct 1;132(4):741–51.
88. Kwak SK, Kim JH. Statistical data preparation: Management of missing values and outliers . Vol. 70, *Korean Journal of Anesthesiology*. *Korean Society of Anesthesiologists*; 2017. p. 407–11.
89. Talge NM, Mudd LM, Sikorskii A, Basso O. United States Birth Weight Reference Corrected For Implausible Gestational Age Estimates. *Pediatrics*. 2014;133(5):844–53.
90. Basso O, Wilcox AJ. Might rare factors account for most of the mortality of preterm babies? *Epidemiology*. 2011;22(3):320–7.
91. Kramer MS, Platt RW, Wen SW, Joseph KS, Allen A, Abrahamowicz M, et al. A New and Improved Population-Based Canadian Reference for Birth Weight for Gestational Age. *Pediatrics*. 2001

- Aug;108(2):e35–e35.
92. Stoltzfus JC. Logistic regression: A brief primer. *Acad Emerg Med*. 2011 Oct;18(10):1099–104.
 93. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012 Sep;24(3):69–71.
 94. Liu J, Tang W, Chen G, Lu Y, Feng C, Tu XM. Correlation and agreement: overview and clarification of competing concepts and measures. *Shanghai Arch psychiatry*. 2016 Apr 25;28(2):115–20.
 95. Derksen S, Keselman HJ. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *Br J Math Stat Psychol*. 1992;45(2):265–82.
 96. Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. *Source Code Biol Med*. 2008 Dec 16;3:17.
 97. Harrell , FE. Regression Modeling Strategies. Cham: Springer International Publishing; 2015. (Springer Series in Statistics). Available from: <http://link.springer.com/10.1007/978-3-319-19425-7>
 98. Austin PC, Steyerberg EW. Interpreting the concordance statistic of a logistic regression model: Relation to the variance and odds ratio of a continuous explanatory variable. *BMC Med Res Methodol*. 2012 Jun 20;12:82.
 99. Hosmer D, Lemeshow S. Applied logistic regression. 2nd ed. New York: John Wiley & Sons; 2000.
 100. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. *Crit Care Med*. 2007 Sep;35(9):2052–6.
 101. Arboretti R, Salmaso L. Model performance analysis and model validation in logistic regression. *Statistica*. 2003 Oct 19;63(2):375–96.
 102. Taylor JMG, Ankerst DP, Andridge RR. Validation of biomarker-based risk prediction models . Vol. 14, *Clinical Cancer Research*. 2008. p. 5977–83.
 103. Fukuda S. Correlation between function of the pituitary-thyroid axis and metabolism of catecholamines by the fetus at delivery. *Clin Endocrinol (Oxf)*. 1987 Sep 1;27(3):331–8.
 104. Lee SY. Perinatal factors associated with neonatal thyroid-stimulating hormone in normal newborns. *Ann Pediatr Endocrinol Metab*. 2016 Dec;21(4):206.
 105. Aktas ON, Gursoy T, Soysal E, Esencan E, Ercin S. Thyroid hormone levels in late preterm, early term and term infants: A study with healthy neonates revealing reference values and factors affecting thyroid hormones. *J Pediatr Endocrinol Metab*. 2017;30(11):1191–6.
 106. Turan S, Bereket A, Angaji M, Koroglu OA, Bilgen H, Onver T, et al. The effect of the mode of delivery on neonatal thyroid function. *J Matern Neonatal Med*. 2007;20(6):473–6.
 107. Schulpis KH, Vlachos GD, Papakonstantinou ED, Karikas GA, Vlachos DG, Roma E, et al. Maternal-neonatal amino acid blood levels in relation to the mode of delivery. *Acta Obstet Gynecol Scand*. 2009;88(1):71–6.

108. Karahan N, Arslan H, Çam Ç. The behaviour of pelvic floor muscles during uterine contractions in spontaneous and oxytocin-induced labour. *J Obstet Gynaecol (Lahore)*. 2018 Jul 4;38(5):629–34.
109. Matsumoto K, Mizuno M, Mizuno T, Dilling-Hansen B, Lahoz A, Bertelsen V, et al. Branched-chain Amino Acids and Arginine Supplementation Attenuates Skeletal Muscle Proteolysis Induced by Moderate Exercise in Young Individuals. *Int J Sports Med*. 2007 Jun;28(6):531–8.
110. McCoin CS, Knotts TA, Adams SH. Acylcarnitines-old actors auditioning for new roles in metabolic physiology . Vol. 11, *Nature Reviews Endocrinology*. NIH Public Access; 2015. p. 617–25.
111. Corr PB, Creer MH, Yamada KA, Saffitz JE, Sobel BE. Prophylaxis of early ventricular fibrillation by inhibition of acylcarnitine accumulation. *J Clin Invest*. 1989 Mar;83(3):927–36.
112. Heathers GP, Yamada KA, Kanter EM, Corr PB. Long-chain acylcarnitines mediate the hypoxia-induced increase in α 1-adrenergic receptors on adult canine myocytes. *Circ Res*. 1987;61(5):735–46.
113. Gucciardi A, Zaramella P, Costa I, Pirillo P, Nardo D, Naturale M, et al. Analysis and interpretation of acylcarnitine profiles in dried blood spot and plasma of preterm and full-term newborns. *Pediatr Res* . 2015 Jan 30 [cited 2018 Sep 10];77(1):36–47.
114. de Boo HA, Harding JE. Protein metabolism in preterm infants with particular reference to intrauterine growth restriction. *Arch Dis Child - Fetal Neonatal Ed*. 2007 Jul;92(4):F315–9.
115. Longini M, Perrone S, Vezzosi P, Marzocchi B, Kenanidis A, Centini G, et al. Association between oxidative stress in pregnancy and preterm premature rupture of membranes. *Clin Biochem*. 2007 Jul;40(11):793–7.
116. Woods JR. Reactive oxygen species and preterm premature rupture of membranes - A review . Vol. 22, *Placenta*. 2001. p. S38–44.
117. Stuart EL, Evans GS, Lin YS, Powers HJ. Reduced Collagen and Ascorbic Acid Concentrations and Increased Proteolytic Susceptibility with Prelabor Fetal Membrane Rupture in Women¹. *Biol Reprod*. 2005 Jan 1;72(1):230–5.
118. Wall PD, Pressman EK, Woods JR. Preterm premature rupture of the membranes and antioxidants: the free radical connection. *J Perinat Med*. 2002 Jan 3;30(6):447–57.
119. Vaisbuch E, Kusanovic JP, Erez O, Mazaki-Tovi S, Gotsch F, Kim CJ, et al. Amniotic fluid fetal hemoglobin in normal pregnancies and pregnancies complicated with preterm labor or prelabor rupture of membranes. *J Matern Neonatal Med*. 2009 May;22(5):388–97.
120. Vaisbuch E, Romero R, Erez O, Kusanovic JP, Gotsch F, Than NG, et al. Total hemoglobin concentration in amniotic fluid is increased in intraamniotic infection/inflammation. *Am J Obstet Gynecol*. 2008 Oct;199(4):426.e1-426.e7.
121. Alger LS, Kisner HJ, Nagey DA. The presence of a meconium-like substance in second-trimester amniotic fluid. *Am J Obstet Gynecol*. 1984 Oct 15;150(4):380–5.
122. Weinberg E. Iron Withholding: A Defense Against Infection and Neoplasia. *Physiol Rev*. 1984 Jan;64(1):65–102.

123. Bank A. Regulation of human fetal hemoglobin: new players, new complexities. *Blood*. 2006 Jan 15;107(2):435–43.
124. Hashimoto F, Nishiumi S, Miyake O, Takeichi H, Chitose M, Ohtsubo H, et al. Metabolomics analysis of umbilical cord blood clarifies changes in saccharides associated with delivery method. *Early Hum Dev*. 2013;89(5):315–20.
125. Guttman A, Vermeulen M, Simeonov D. Are rates of prenatal ultrasound a valid measure of health system overuse? 2012.
126. Stout MJ, Busam R, Macones GA, Tuuli MG. Spontaneous and indicated preterm birth subtypes: Interobserver agreement and accuracy of classification. *American Journal of Obstetrics and Gynecology*. NIH Public Access; 2014. p. 530.e1-530.e4.

APPENDIX A: Certificates for Ethical Approval

A.I CHEO Research Ethics Board Certificate



FW: REB Protocol No 17/139X - Final Approval - Delegated Review

Hi everyone,

Just confirmation that the student project for Jessica Yau has received REB approval from CHEO.

Best,

Deshayne

Subject: REB Protocol No 17/139X - Final Approval - Delegated Review



RESEARCH INSTITUTE
INSTITUT DE RECHERCHE

**CHEO Research Ethics Board
Approval - Delegated Review**

Approval Date*: August 11, 2017

Valid Until:** July 15, 2018

Annual Renewal Submission Deadline: June 15, 2018

Documents Reviewed & Approved:

Document Name	Comments	Version Date
Protocol	Protocol	2017/08/09

This is to notify you that the Children's Hospital of Eastern Ontario Research Ethics Board has granted approval to the above named research study on the date noted above. Your project was reviewed under the delegated review stream, which is reserved for projects that involve no more than minimal risk to human subjects.

Final approval is granted for the above noted study, with the understanding that the investigator agrees to comply with the following requirements:

1. The investigator must conduct the study in compliance with the protocol and any additional conditions set out by the Board.
2. The investigator is responsible for complying with all applicable guidelines and regulations regarding human research ethics conduct, as applicable to the research project.
3. Approval for studies that include an investigational device(s) is contingent upon the investigator securing an Investigational Testing Authorization notice from Health Canada.
4. Investigators must submit an annual renewal report to the REB 30 days prior to the expiration date stated above.
5. The investigator must not implement any deviation from, or changes to, the protocol, consents or assents without the approval of the REB.
6. The investigator must, prior to use, submit to the Board changes to the study documentation, e.g., changes to the informed consent letters, recruitment materials.
7. Investigators must provide the Board with French versions of the consent form, unless a waiver has been granted. An interpreter should be offered to participants as required or at the request of the participant throughout the course of research.
8. The investigator must promptly report to the REB all unexpected and untoward occurrences (including the loss or theft of study data and other such privacy breaches).
9. Investigators must notify the REB of any study closures (closed to accrual, temporary, premature or permanent).
10. Investigators must submit a final report at the conclusion of the study.

Should you have any questions or concerns, please do not hesitate to contact the Research Ethics Board Office at 613-737-7600 ext. 3350 or 2128.

Regards,

Richard Carpentier, PhD

Chair, Research Ethics Board

Président, Comité d'éthique de la recherche

*The final approval date for initial delegated study applications approved with or without modifications will be the date the REB has determined that the conditions of approval have been satisfied.

**The expiry date of REB approval for initial study application that required no modifications will be as follows:

- If the date of review and approval was **on or before** the 15th of the month, the expiry date will be the 15th of the month prior to the date of review and approval by the Chair and/or delegate *in the following year*;
- If the date of review and approval was **after** the 15th the expiry date will be the 15th of the month in which the date of review and approval by the REB *in the following year*.

The expiry date of REB approval for initial study applications that **require modifications** will be as follows:

- If the initial feedback was sent **on or before** the 15th of the month, the expiry date will be the 15th of the month prior to the date the letter of REB feedback is issued to the investigator(s) *in the following year*;

7/11/2018

University of Ottawa | Université d'Ottawa Mail - FW: REB Protocol No 17/139X - Final Approval - Delegated Review

- If the initial feedback was sent after the 15th the expiry date will be the 15th of the month in which the feedback was sent *in the following year*.

A.2 CHEO Research Ethics Board Renewal Certificate

7/10/2018

University of Ottawa | Université d'Ottawa Mail - FW: REB Protocol No: 17/139X - Annual Renewal Certificate 2018



FW: REB Protocol No: 17/139X - Annual Renewal Certificate 2018

Wed, Jun 20, 2018 at 8:48 AM

Subject: REB Protocol No: 17/139X - Annual Renewal Certificate 2018

Research Ethics Board 2018 Annual Renewal Approval Letter

Principal Investigator: Ms. Deshayne Fell

REB Protocol No: 17/139X

Romeo File No: 20170399

Project Title: CHEOREB# 17/139X - Association of newborn screening analytes with type of delivery among preterm and term births

Primary Affiliation: Clinical Research\Epidemiology

Protocol Status: Active

Approval Date: June 19, 2018

Approval Expiry Date: July 15, 2019

This is to notify you that the CHEO REB has granted approval to the renewal for the above named research study for a period of one year. The renewal was reviewed in the delegated stream and approved by the Chair or a delegate of the Chair. Decisions made by the Chair under delegated review are ratified by the full Board at its subsequent meeting.

Approval is granted with the understanding that the investigator agrees to comply with the following requirements:

1. The investigator must conduct the study in compliance with the protocol and any additional conditions set out by the Board.
2. The investigator is responsible for complying with all applicable guidelines and regulations regarding the ethical conduct of research with humans, as applicable to the research project.
3. Investigators must obtain annual renewal approval prior to the expiry date stated above.
4. The investigator must not implement any deviation from, or changes to, the protocol without the approval of the REB except where necessary to eliminate an immediate hazard to the research subject, or when the change involves only logistical or administrative aspects of the study (e.g., change of telephone number or research staff).


https://mail.google.com/mail/u/1/?ui=2&ik=93e5f4f331&jsver=udqAzWhC2a4.en.&cbl=gmail_fe_180701.15_p4&view-pt&msg=1641d3b2c0fb8930&se... 1/2

As soon as possible, however, the implemented deviation or change, the reasons for it and, if appropriate, the proposed protocol amendment(s) should be submitted to the Board for review and approval.

5. The investigator must, prior to use, obtain approval from the Board for changes to the study documentation, e.g., changes to the informed consent letters, recruitment materials.
6. Investigators must obtain approval from the Board of French version(s) of the consent/assent form(s), unless a waiver has been granted. An interpreter should be offered to participants as required or at the request of the participant throughout the course of research.
7. For clinical drug or device trials, investigators must promptly report to the REB all adverse events that are both serious and unexpected (SAEs) or unexpected and untoward occurrences (including the loss or theft of study data and other such privacy breaches).
8. For SAE reports on clinical drug trials, the investigator must also comply with the hospital-wide Policy regarding Procedures for Considering Medical Error in the Differential Diagnosis of Severe Adverse Events (SAE) Associated with the Drugs Administered in a Clinical Trial.
9. Investigators must promptly report to the REB any new information regarding the safety of research subjects (e.g., changes to the product monograph or investigator's brochure of drug trials). Where available, any reports produced by the Data Safety Monitoring Board should also be promptly submitted to the REB for acknowledgement.
10. Investigators must notify the REB of any study closures (closed to accrual, temporary, premature or permanent).
11. Investigators must submit a study closure event form at the conclusion of the study.

If you have any questions, pertaining to this letter, please contact the Research Ethics Board Office at (613) 737-7600, ext. 3350 or 2128.

Regards,


Richard Carpentier, PhD
Chair, Research Ethics Board
Président, Comité d'éthique de la recherche

|

This message, including any attachments, may contain confidential information and is for the sole use of the intended recipient(s). Any unauthorized use, disclosure or distribution is prohibited. If you are not the intended recipient, please notify the sender immediately and destroy the original message (for further detail please see <http://www.cheo.on.ca/en/disclaimer>). Ce message, y compris les pièces jointes, peut contenir des renseignements confidentiels, et seuls les destinataires visés peuvent le consulter. Il est strictement interdit de l'utiliser sans autorisation, de le divulguer ou de le distribuer. Si ce message ne vous était pas destiné, veuillez en informer l'expéditeur immédiatement et détruire le message original (Veuillez consultez <http://www.cheo.on.ca/avis-non-responsabilite> pour plus de précisions).

APPENDIX B: Letters of Support

BORN Letter of support



Early health. Lifelong health.
Début en santé. Longue vie en santé.

May 30, 2017

Dr. Carpentier
Chair – CHEO REB
401 Smyth Rd., Ottawa, ON
K1H 8L1

Dear Dr. Carpentier:

I am writing to confirm that BORN Ontario has reviewed the protocol titled ***Association of newborn screening analytes with type of delivery among preterm and term births*** submitted by Dr. Deshayne Fell and colleagues.

As per our procedure, we are now writing to you to confirm that BORN is supportive of this project and its scientific value, that proper procedures will be followed and that we are confident that the PHI will be protected. We can also confirm that BORN has a signed Data Sharing Agreement with ICES and that there is a standardized process in place between BORN and ICES for reviewing and approving projects using the linked data.

If you have further questions specifically about the project, they can be directed to Dr. Fell. If you have any questions about the use of BORN data elements, please direct these to myself.

Sincerely,

Sandra Dunn, RN, PhD
Knowledge Translation Specialist, BORN Ontario

BORN: Better Outcomes Registry & Network
Registre et Réseau des Bons Résultats dès la naissance

BORN Ontario | CHEO Research Institute | Centre for Practice-Changing Research Building
401 Smyth Road | Ottawa, ON | K1H 8L1 | T: 613.737.7600 x 6022 | F: 613.737. 6504 | www.BORNOntario.ca

NSO Letter of Support



April 3rd 2017

RE: Use of NSO data at ICES

To whom it may concern,

As the Newborn Screening Ontario (NSO) data custodian, I am writing to provide my permission for use of NSO data for the projects listed in the attached table.

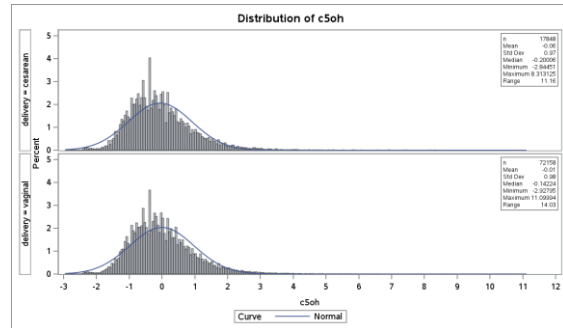
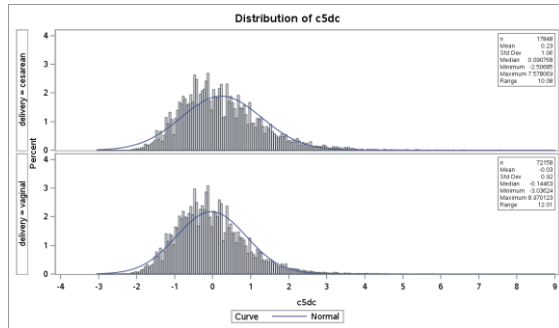
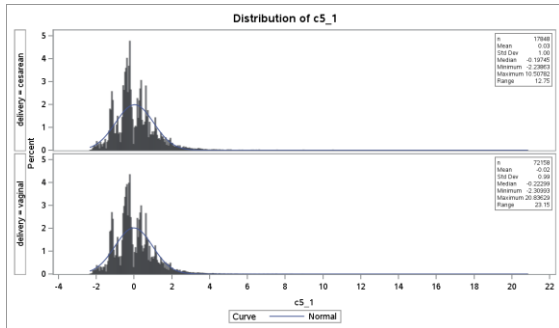
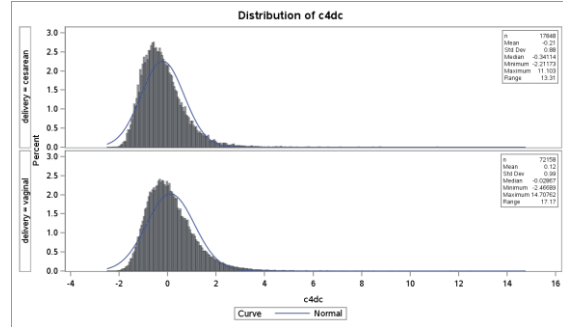
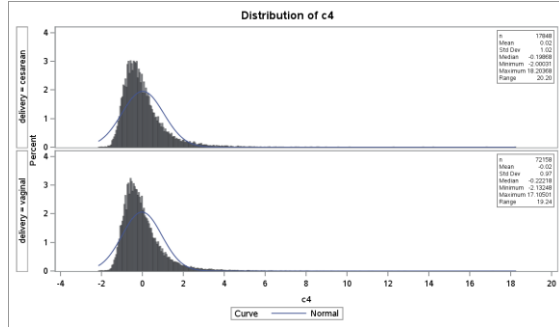
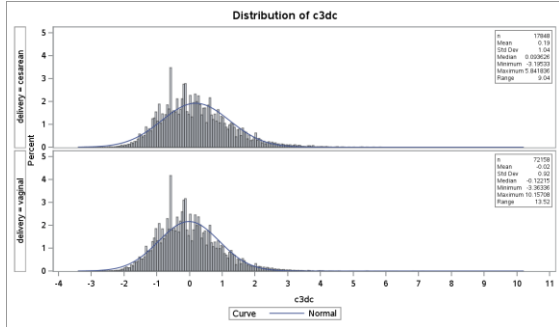
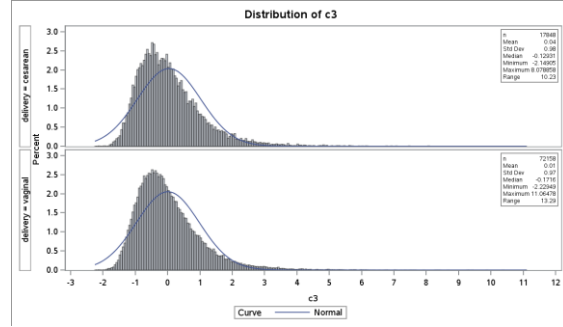
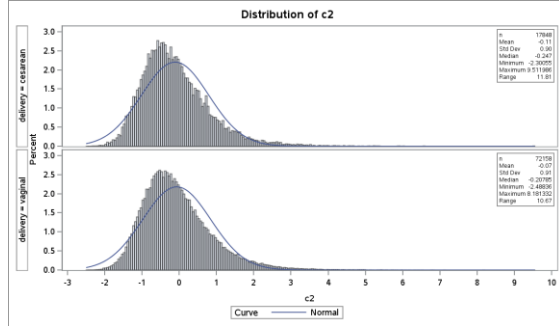
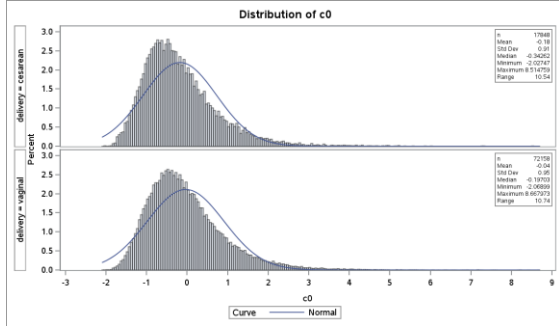
Sincerely,

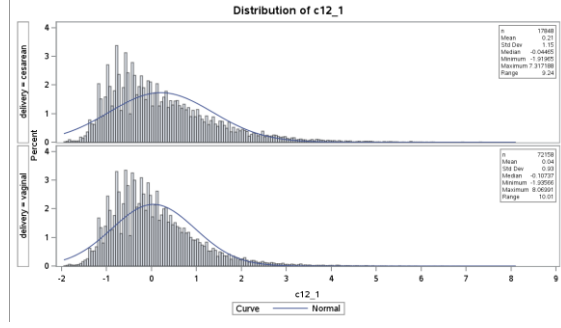
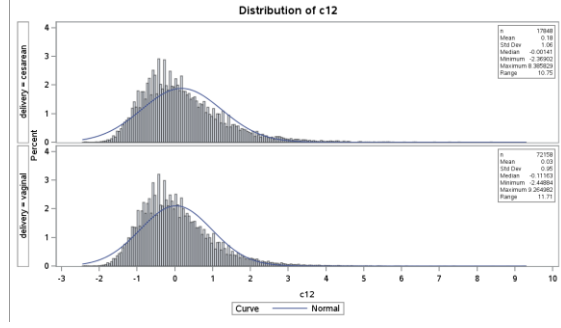
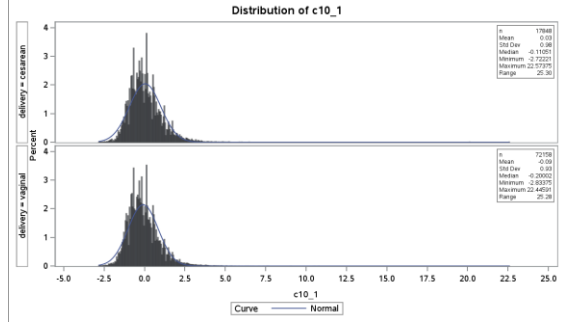
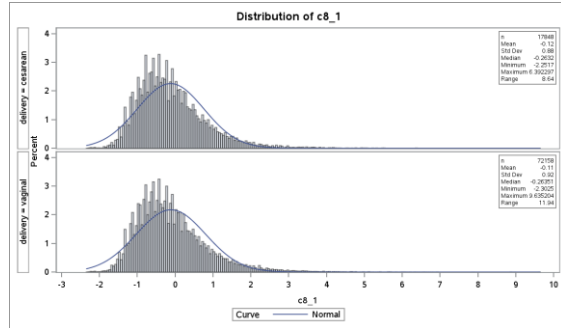
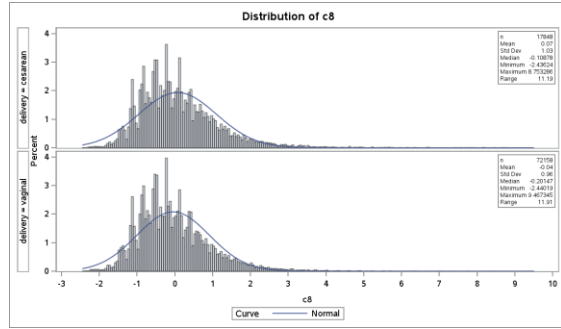
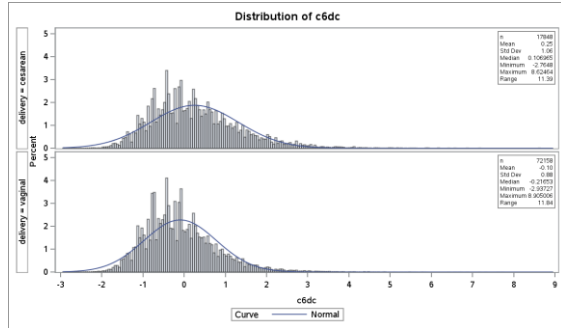
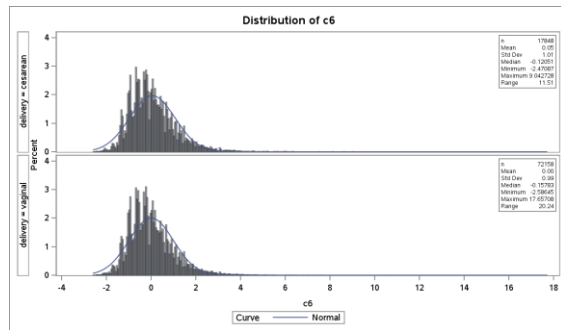
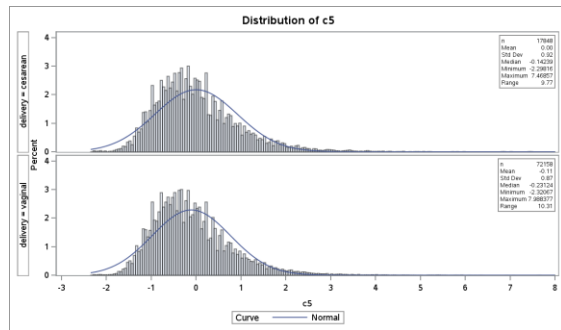
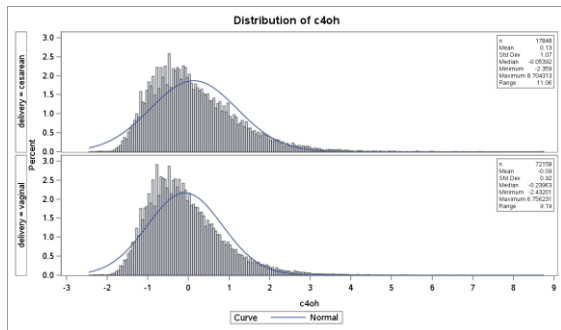
Pranesh Chakraborty, MD, FRCPC, FCCMG
Physician, Section of Metabolism and Newborn Screening, Department of Pediatrics, CHEO
Medical Director, BORN Ontario
Associate Professor, Department of Pediatrics, Faculty of Medicine, University of Ottawa

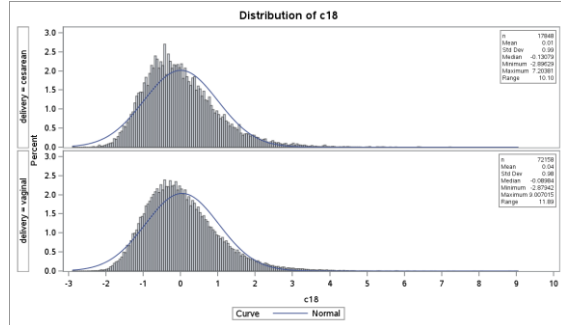
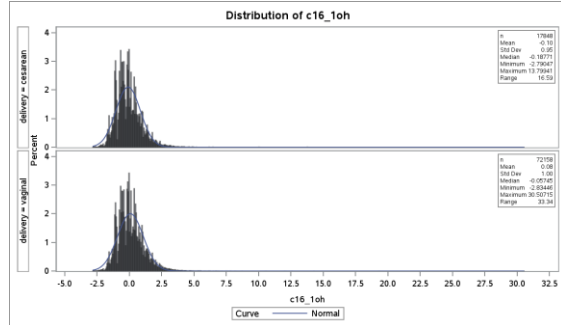
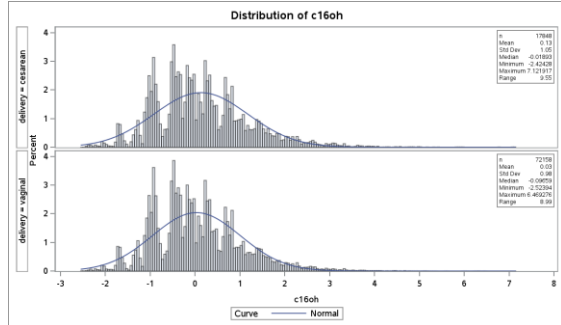
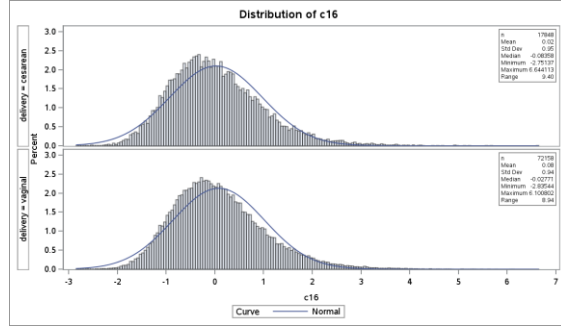
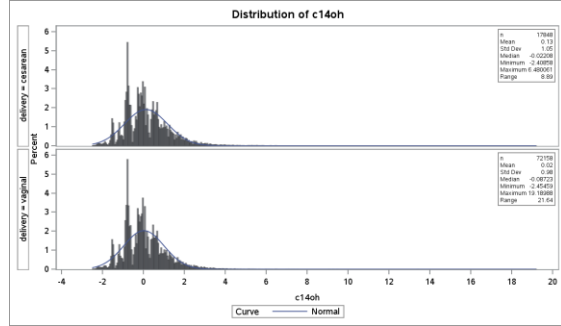
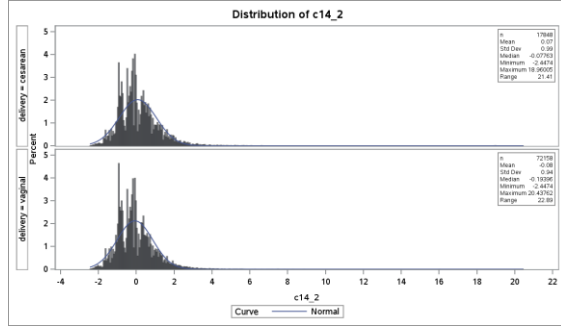
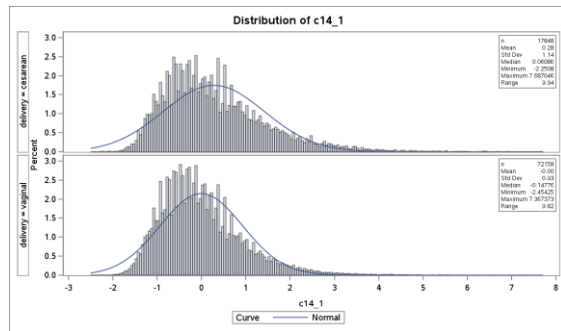
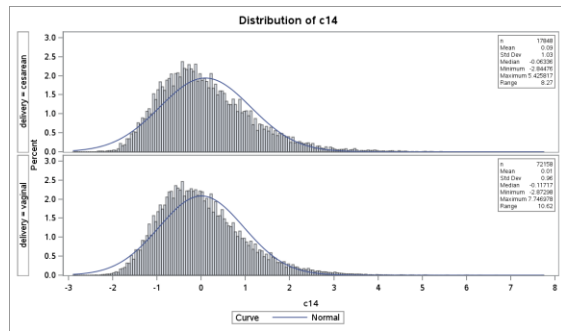
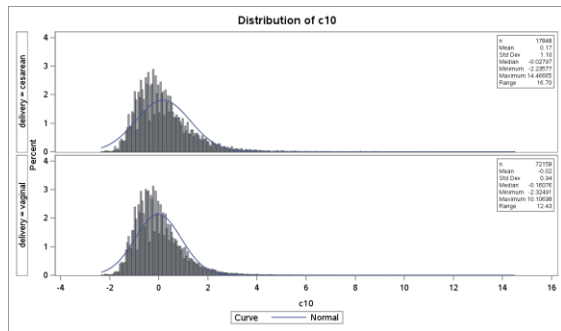
NSO data uses at ICES

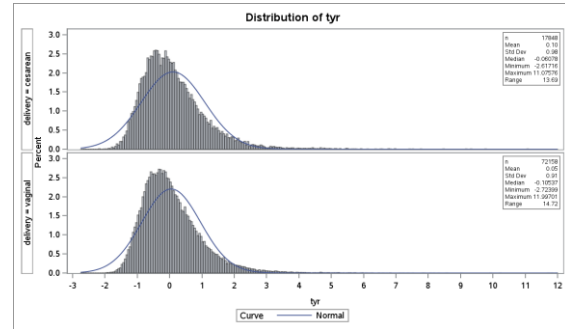
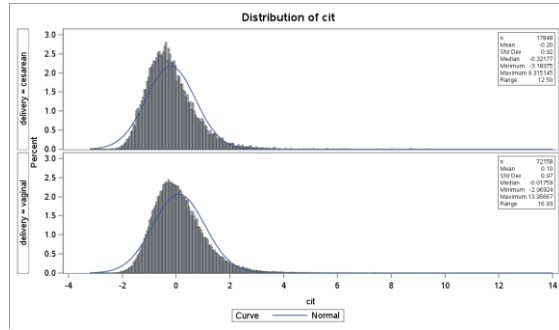
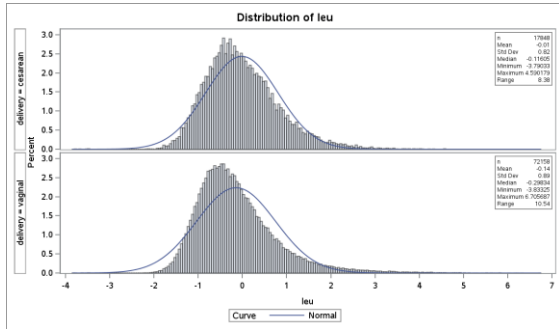
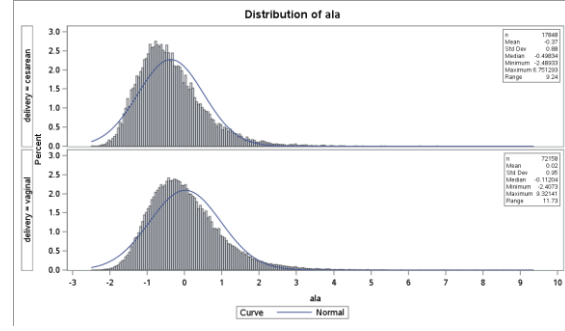
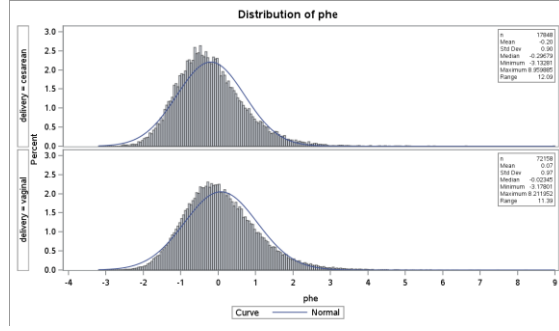
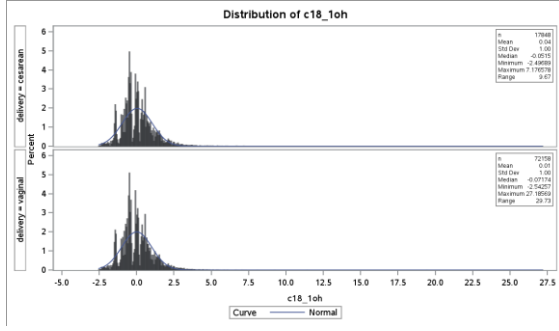
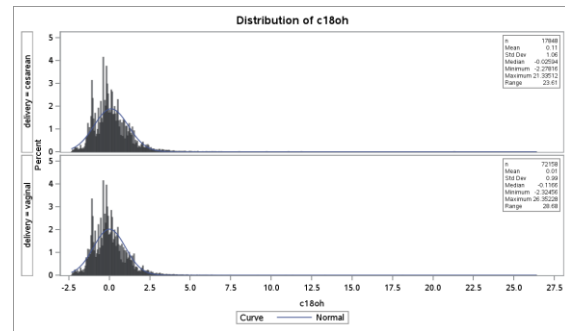
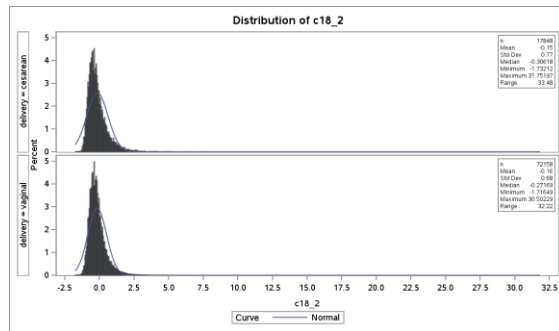
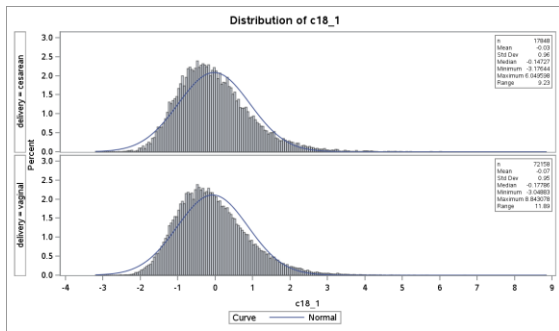
Research project title	Analysis title	Lead	TRIM number	Status										
Investigating the associations between newborn screening results and specific health outcomes, and using these to build predictive models	Predicting Future Health Status Based on Newborn Screening Metabolite Levels and Other Markers of Perinatal Health in the General Population and in Vulnerable Subgroups of Children Born in Ontario	Kumanan Wilson (OHRI)	2015 0901 065 000	Analysis completed, results dissemination ongoing										
	Incremental predictive value of metabolic biomarkers at birth and the development of chronic kidney disease or kidney failure in early childhood	Kumanan Wilson and Manish Sood (OHRI)	2016 0901 094 000	Analysis completed, results dissemination ongoing										
	Predicting Gestational Age Using Newborn Screening Data	Kumanan Wilson/Steven Hawken OHRI	TBD	Analysis completed, results dissemination ongoing										
Descriptive epidemiology and health services impact of inborn errors of metabolism in an Ontario newborn cohort	Investigation of newborn screening markers and other birth characteristics for the prediction of neonatal sepsis, birth	Deshayne Fell/Steven Hawken/Kumanan Wilson OHRI	2017 0901 115 000	Analysis completed, results dissemination ongoing										
	Metabolomics of Birth: Assessing the influence of environmental and physiological factors on infant metabolic profiles	Steven Hawken/Kumanan Wilson (OHRI)	2017 0901 125 000	Analysis ongoing?										
Descriptive epidemiology and health services impact of MCADD in an Ontario newborn cohort	Descriptive epidemiology and health services impact of PKU in an Ontario newborn cohort	Pranesh Chakraborty/Beth Potter, uOttawa/CHEO	TBD	Analysis completed, results dissemination ongoing										
	Descriptive epidemiology and health services impact of MCADD in an Ontario newborn cohort	Pranesh Chakraborty/Beth Potter, uOttawa/CHEO	TBD	Analysis completed, results dissemination ongoing										
Descriptive epidemiology of haemoglobinopathies in an Ontario newborn cohort: Results from the first three years of screening at Newborn Screening Ontario Association of newborn screening analytes with type of delivery among term and preterm births	"	Beth Potter, uOttawa	TBD	Analyses completed, manuscript in development										
	"	Deshayne Fell, CHEO/OHRI	TBD	In development										

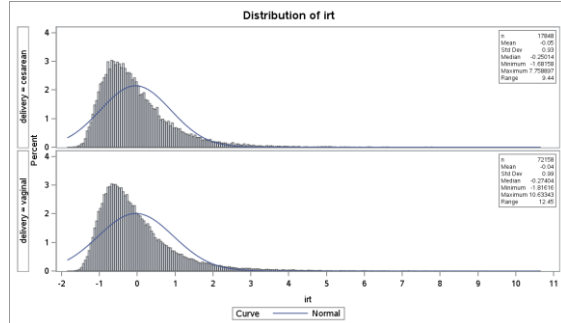
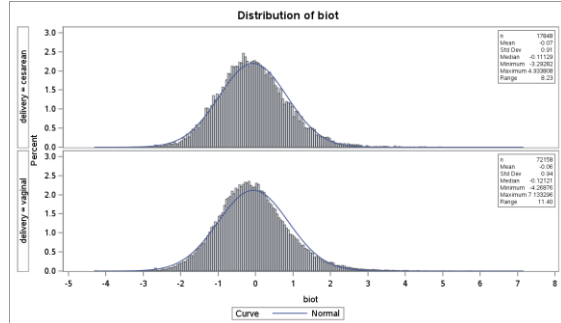
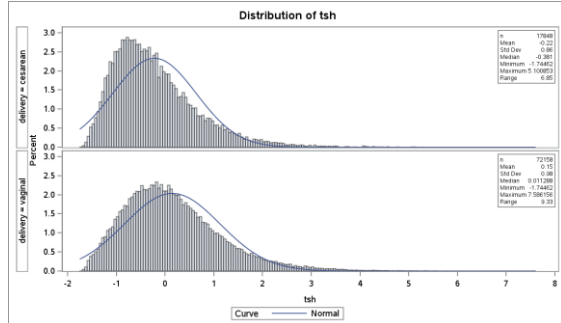
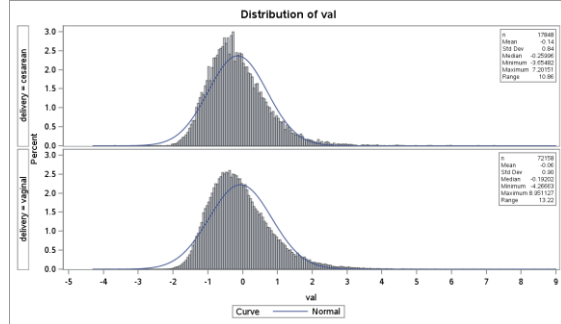
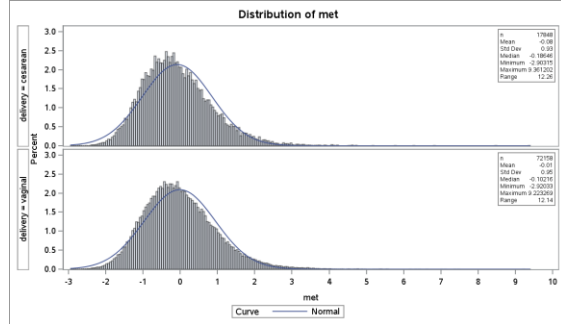
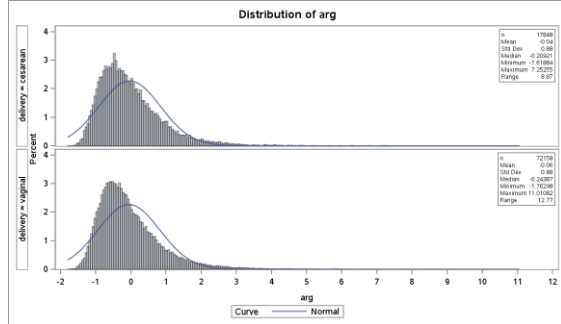
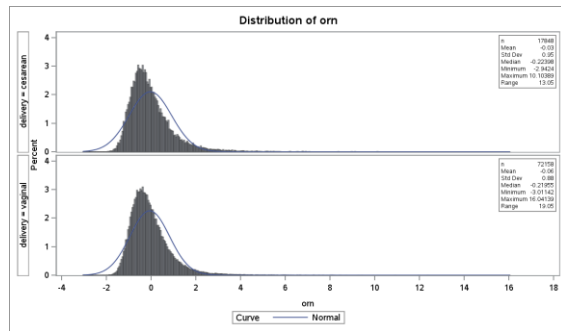
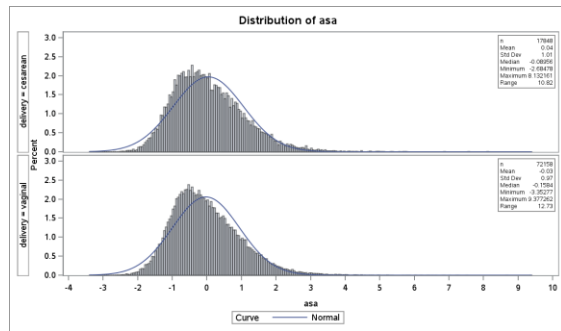
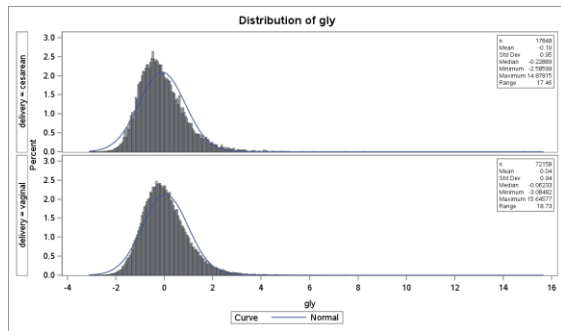
APPENDIX C: Histograms for Objective 1

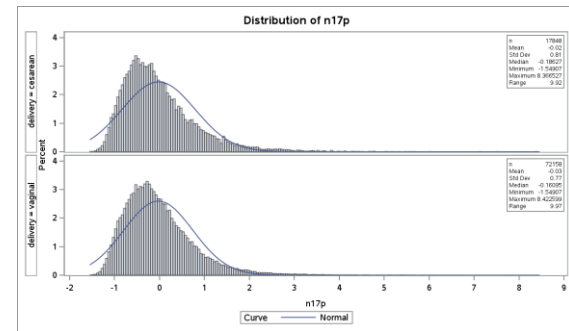
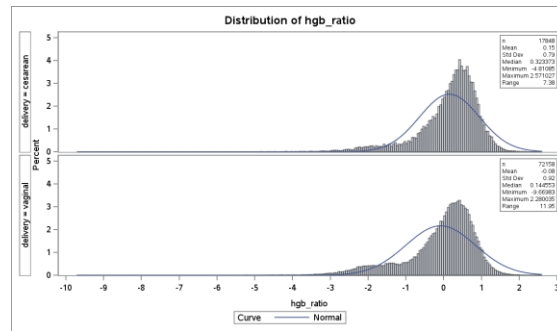
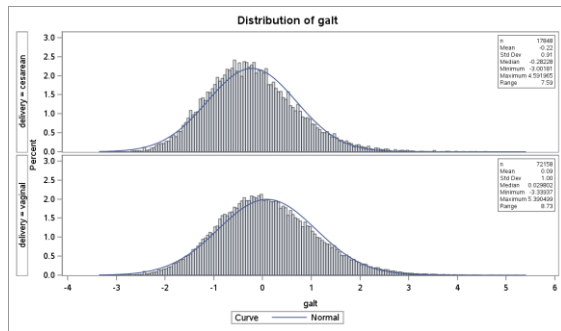




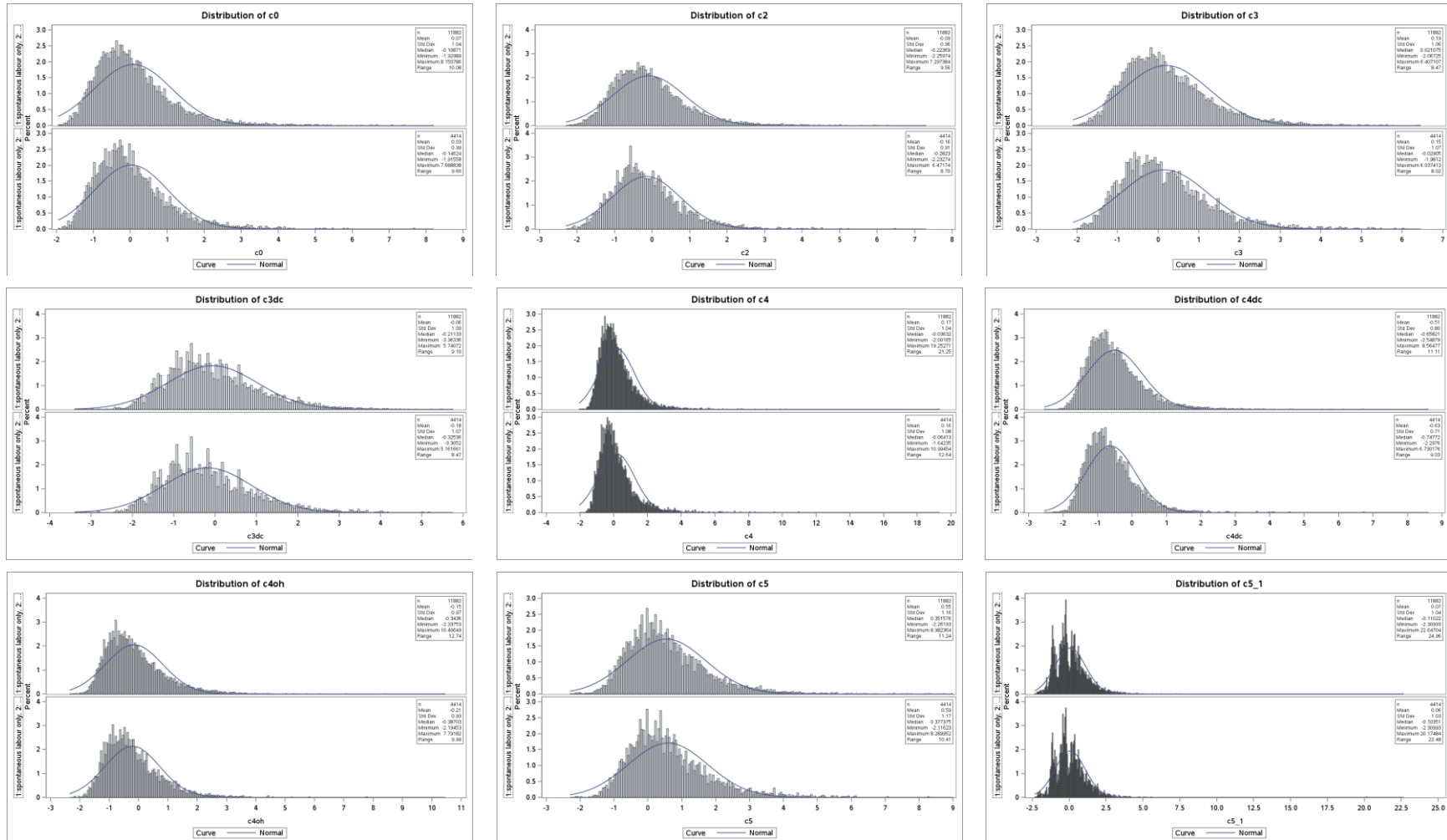


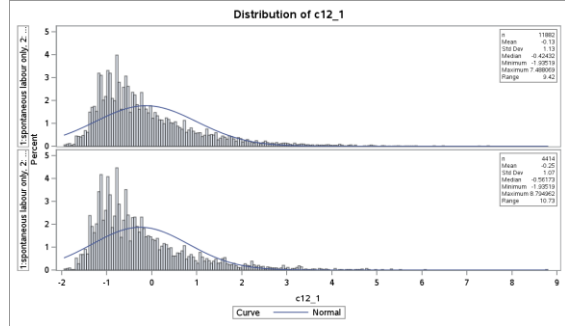
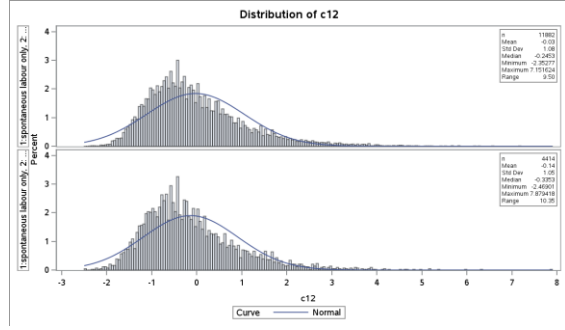
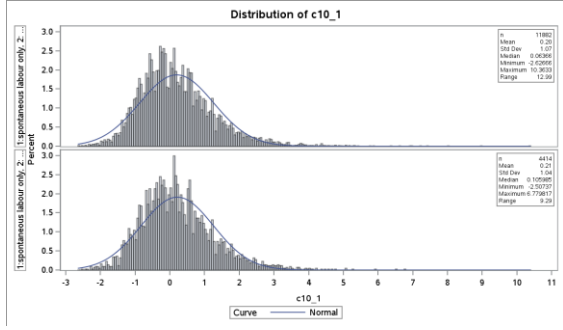
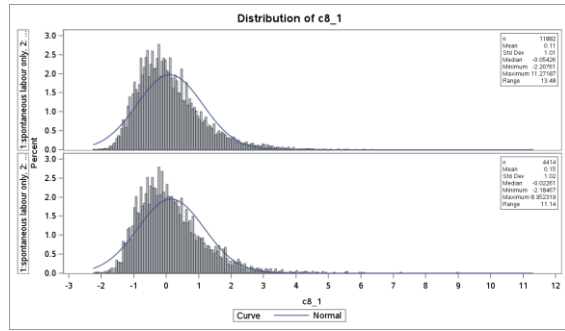
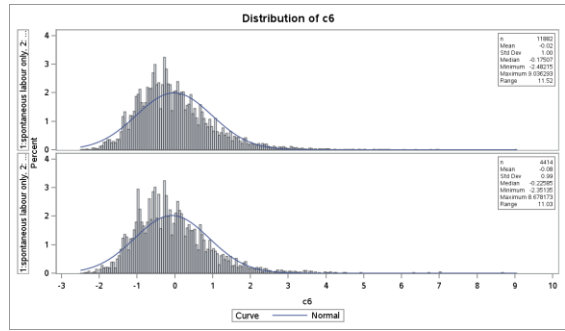
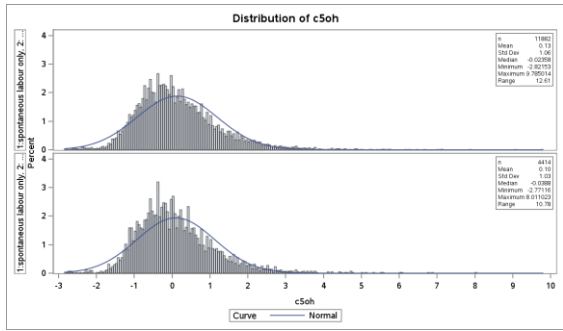
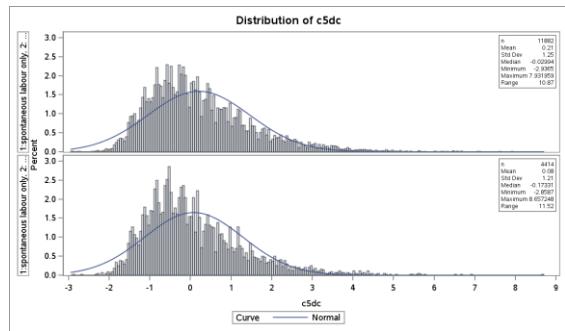
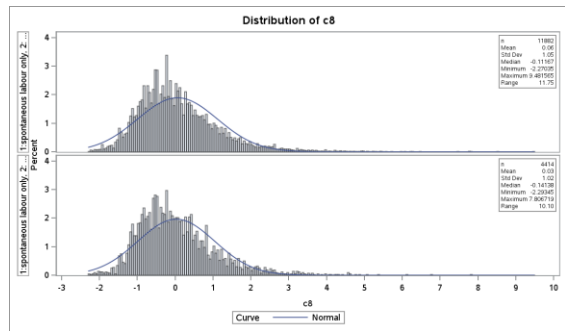
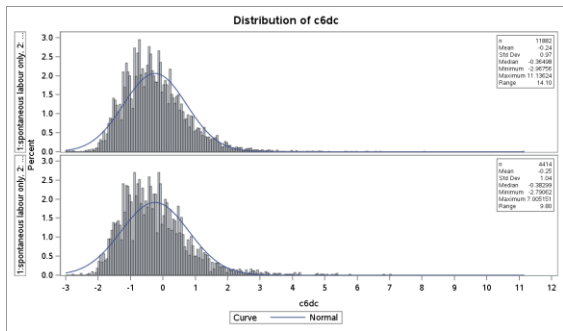


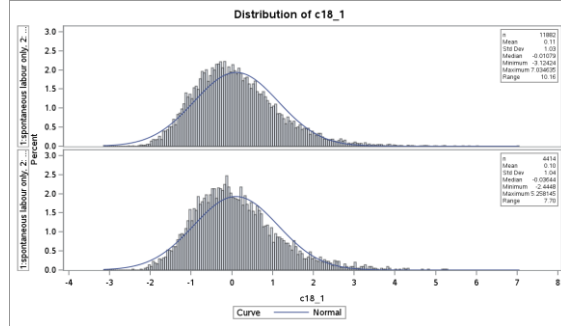
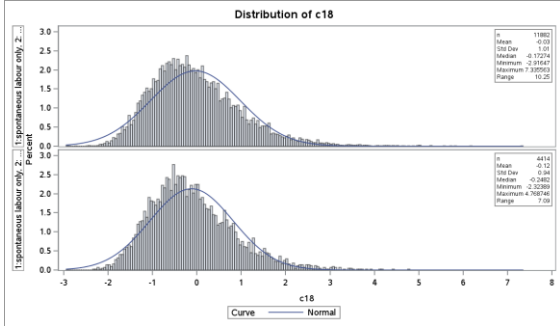
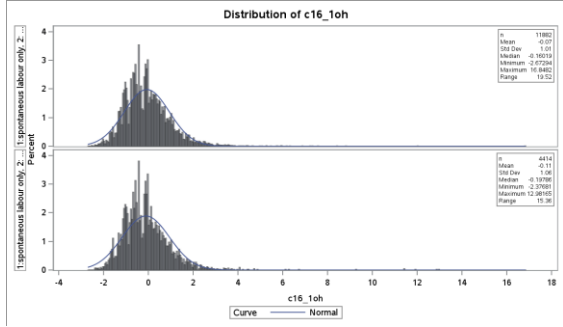
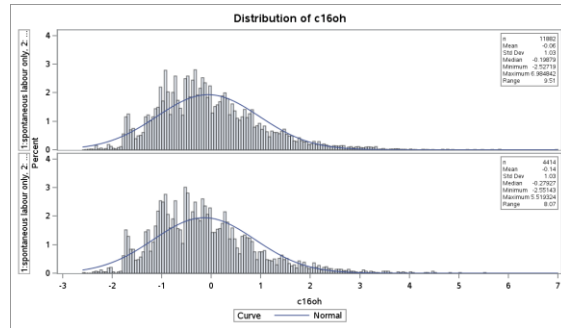
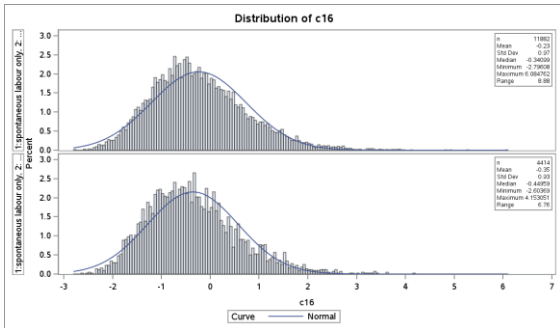
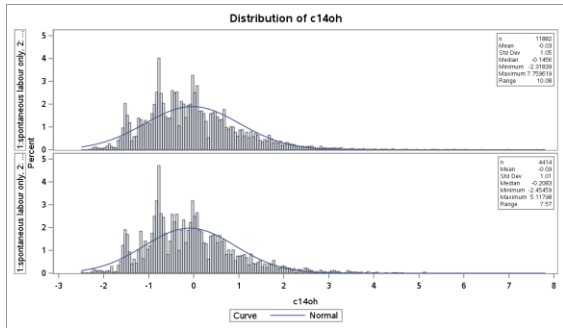
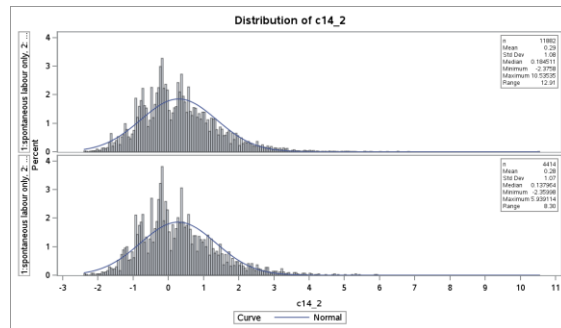
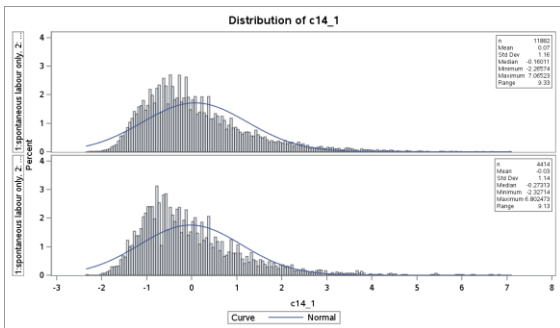
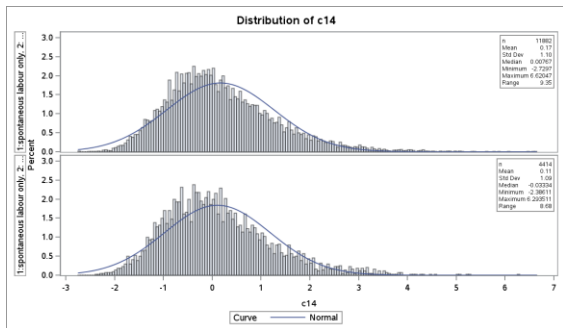


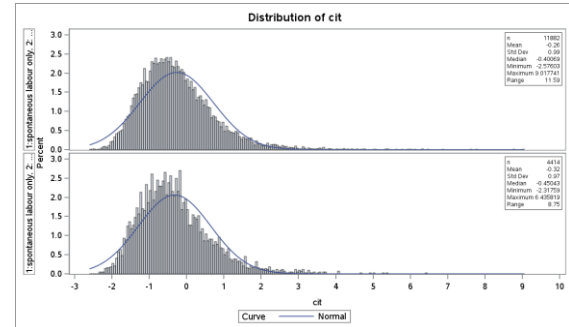
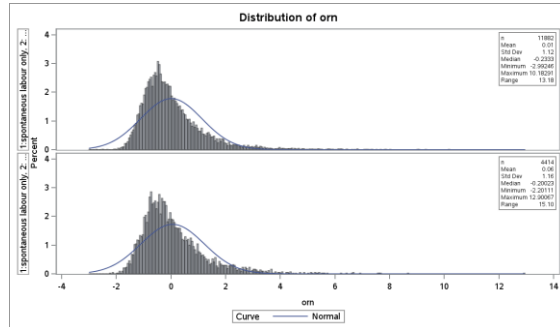
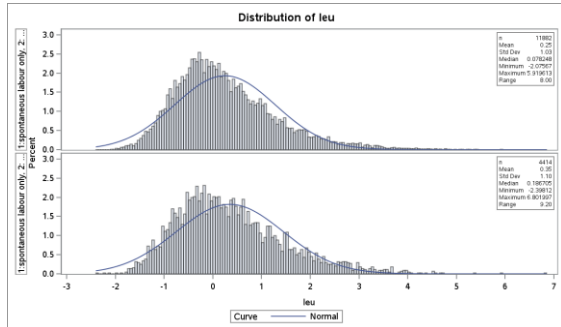
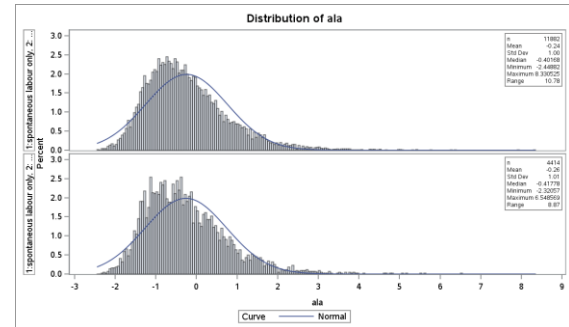
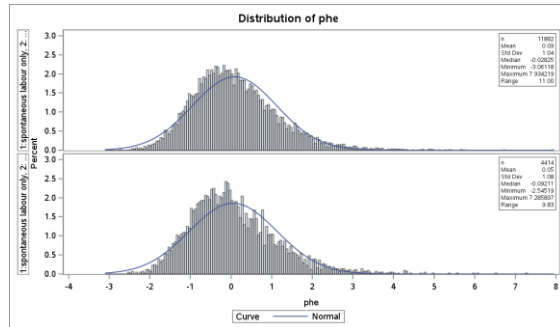
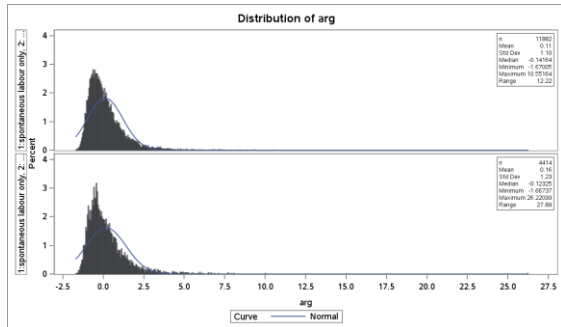
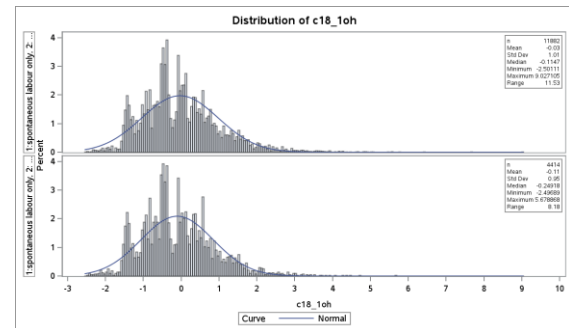
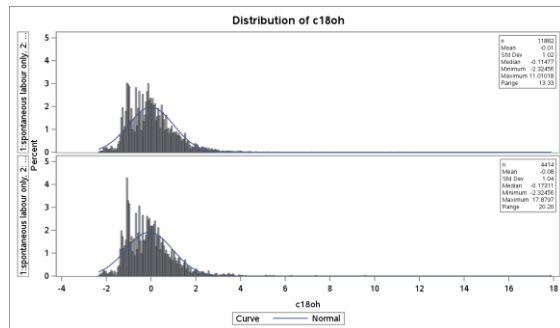
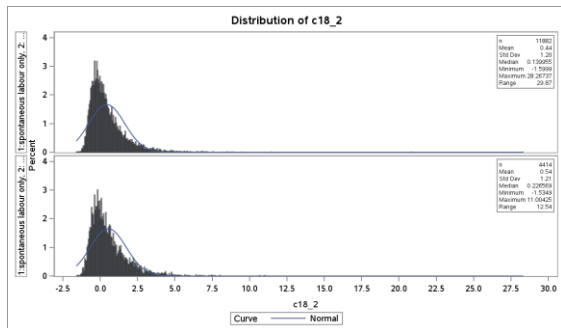


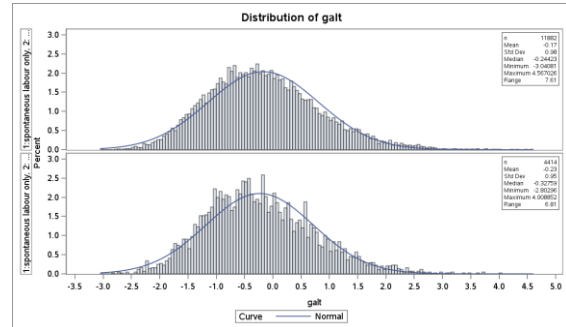
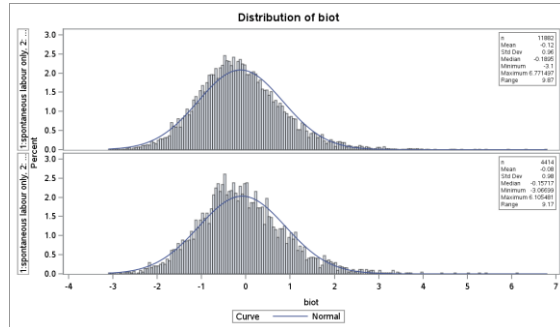
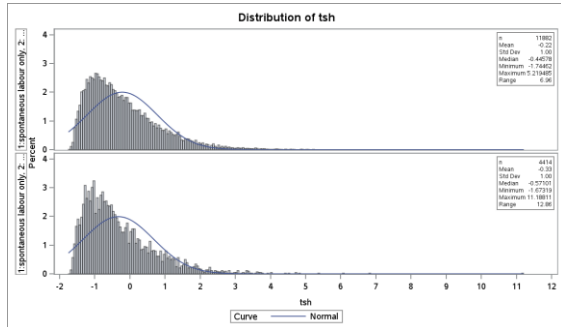
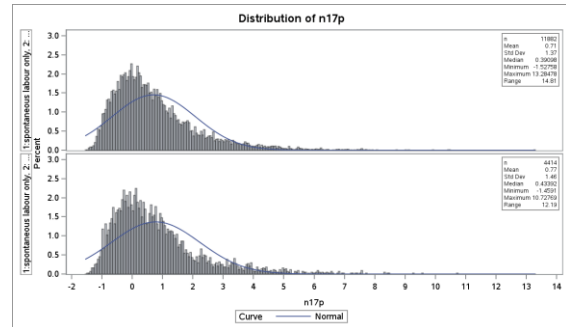
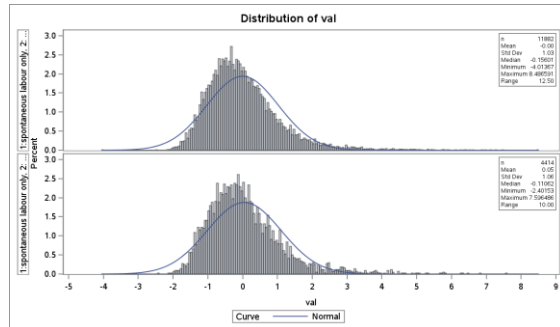
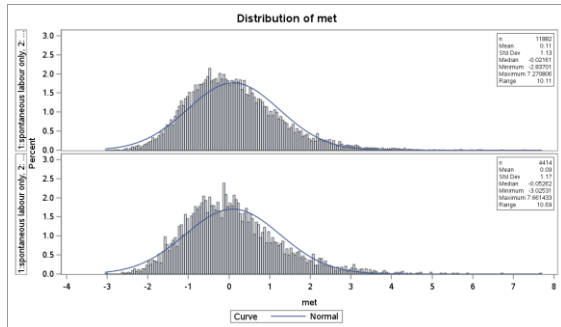
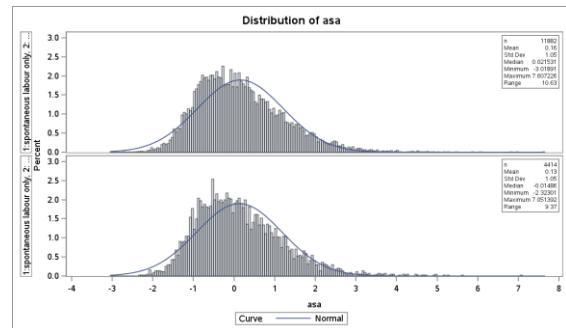
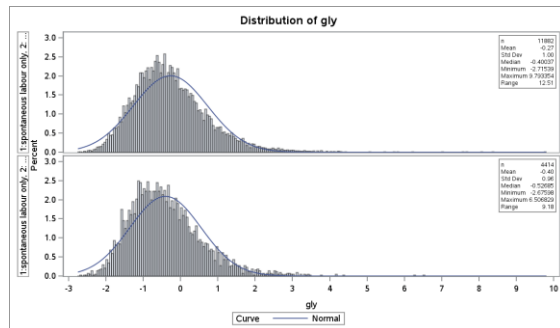
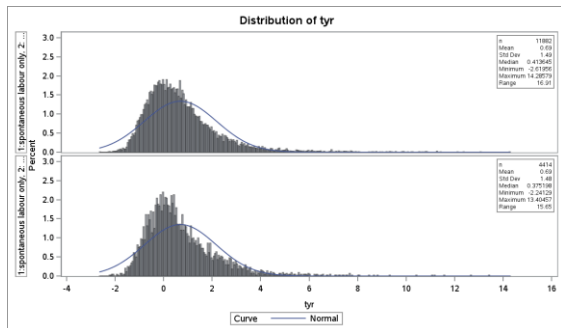
APPENDIX D: Histograms for Objective 2

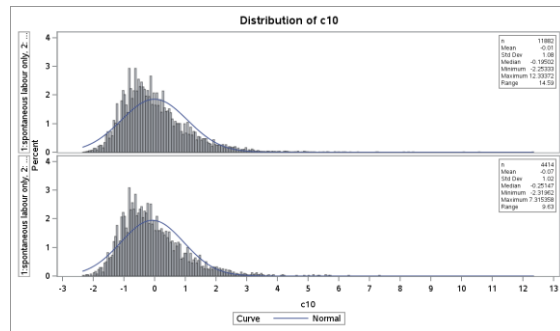
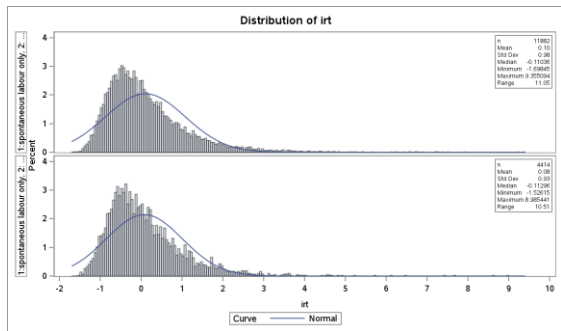




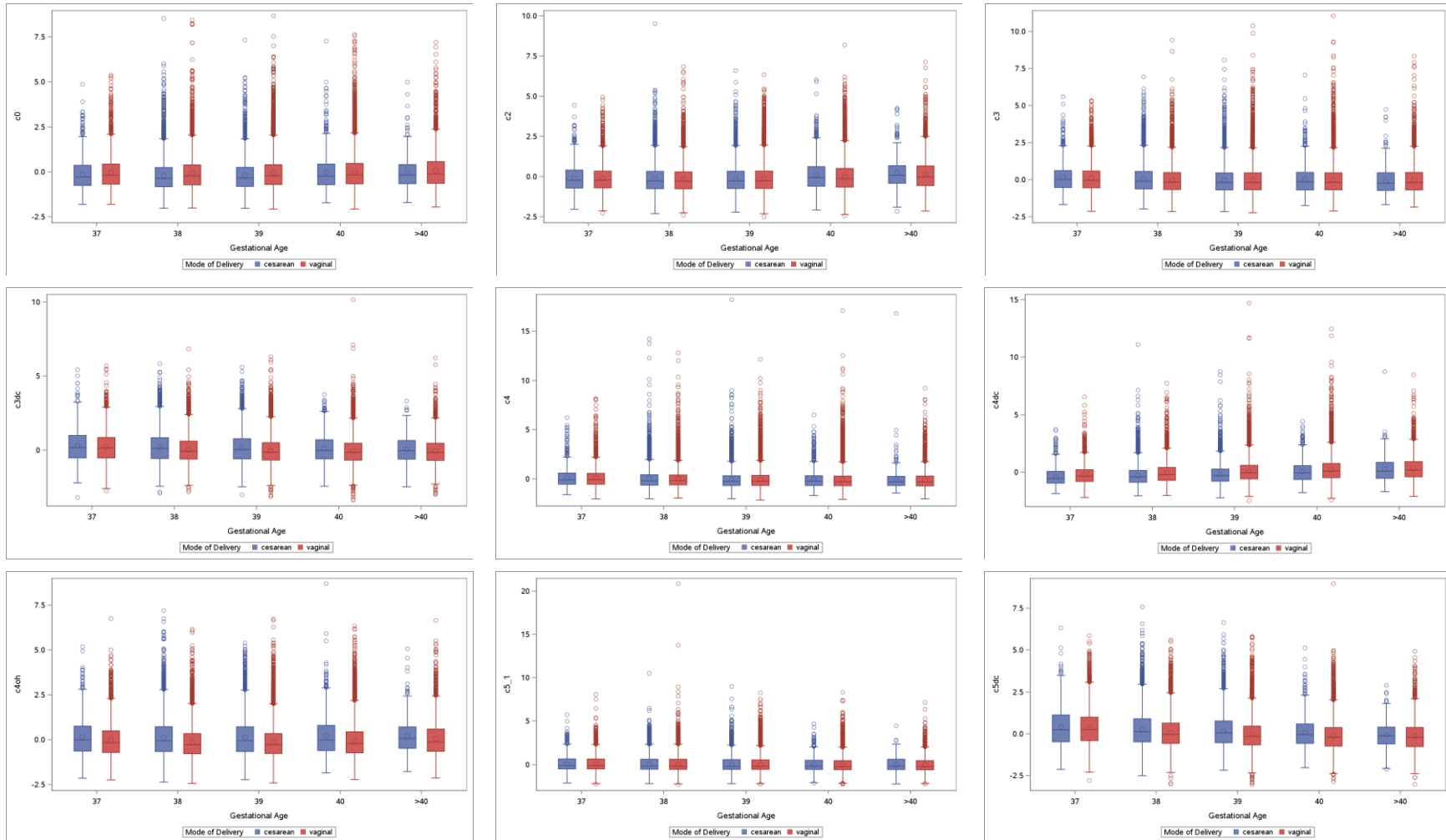


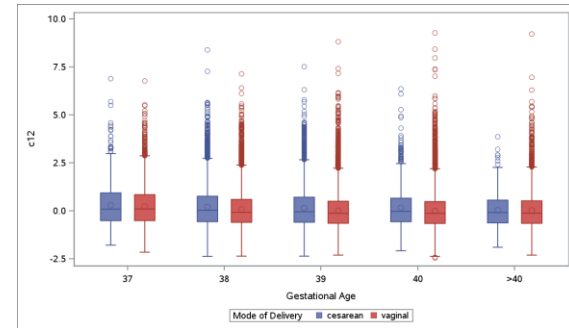
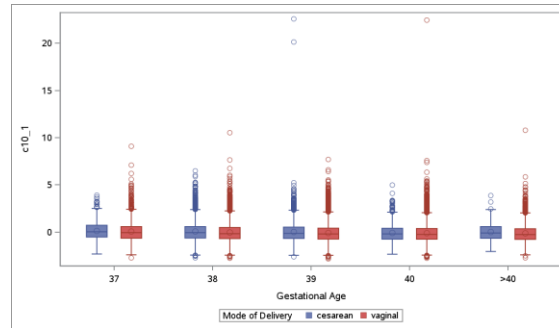
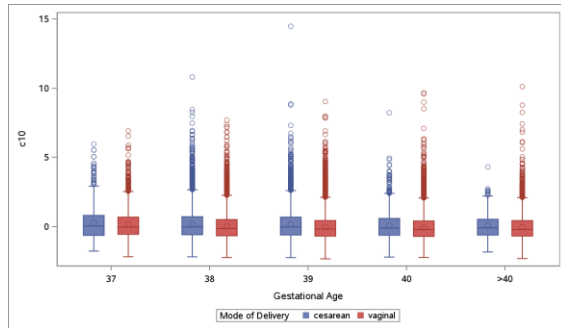
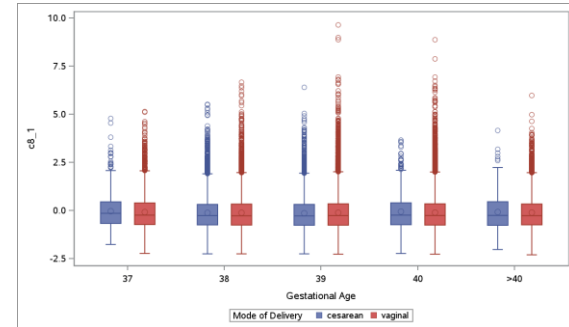
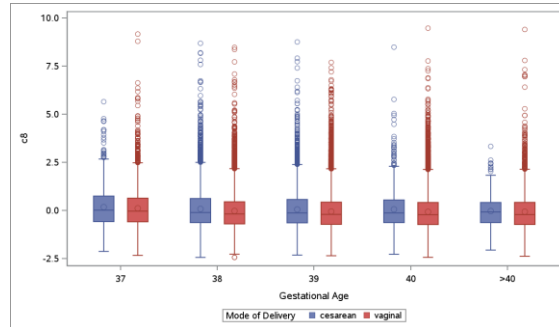
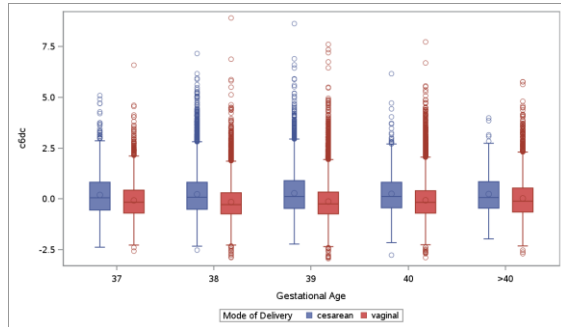
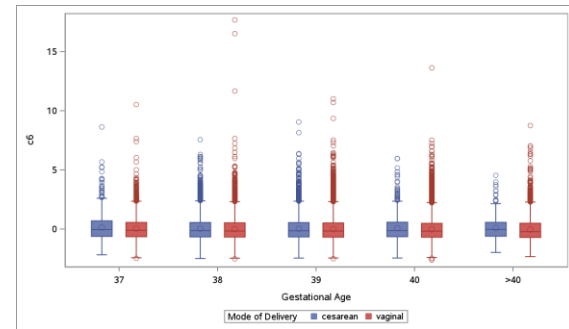
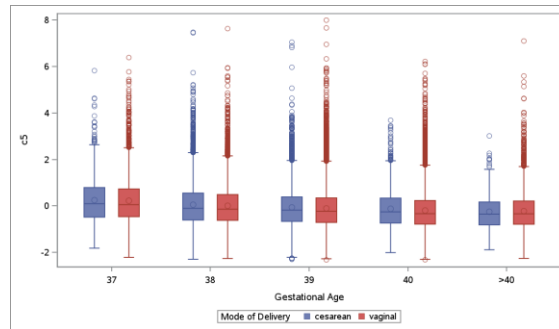
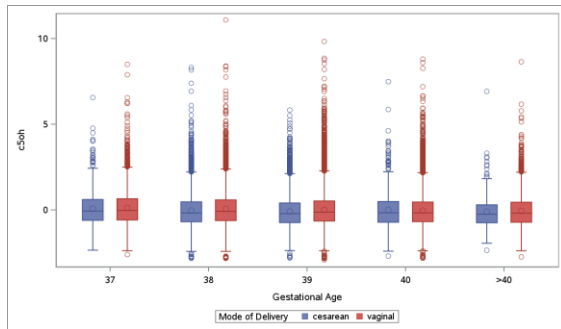


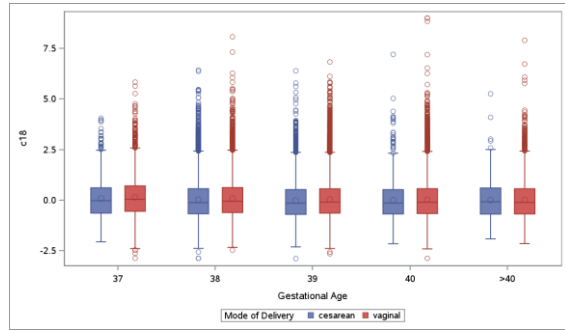
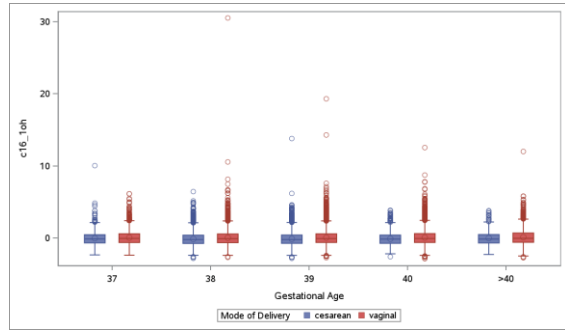
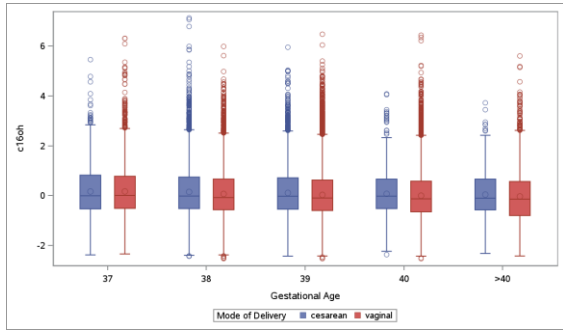
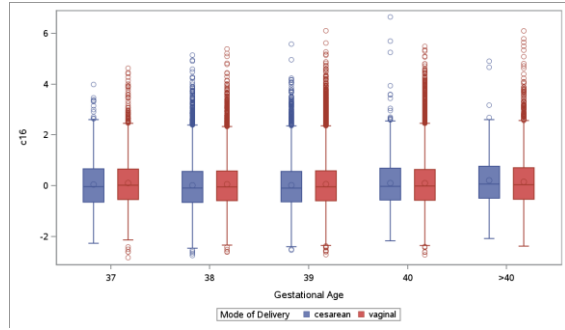
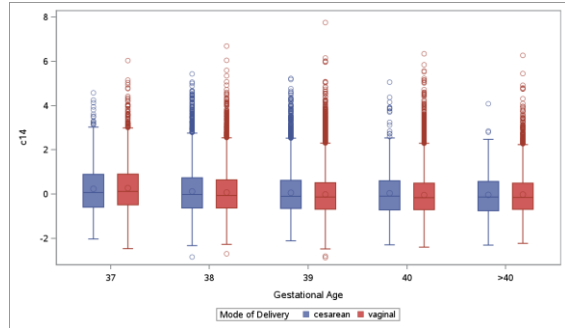
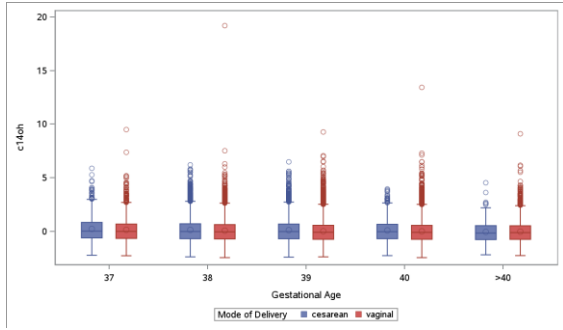
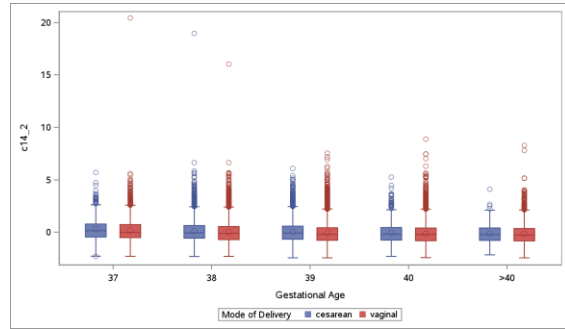
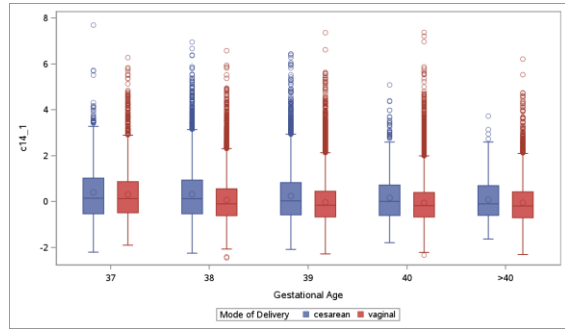
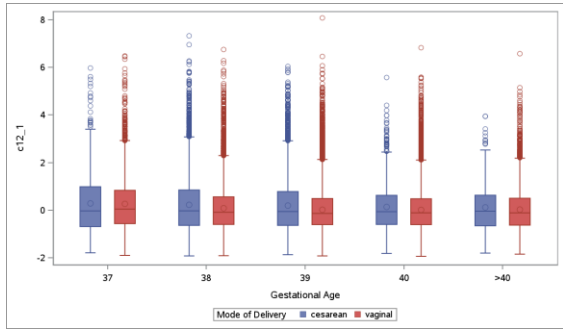


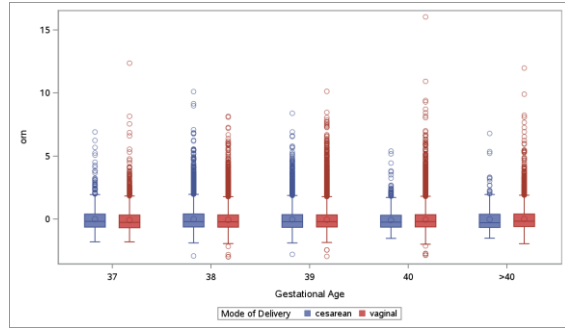
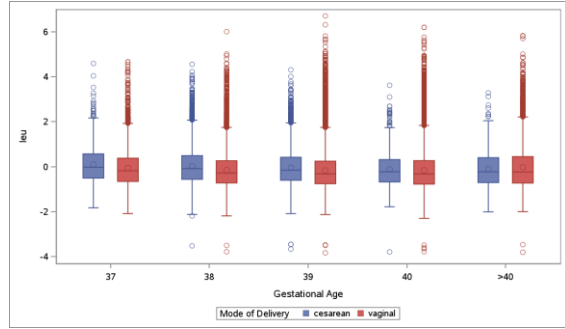
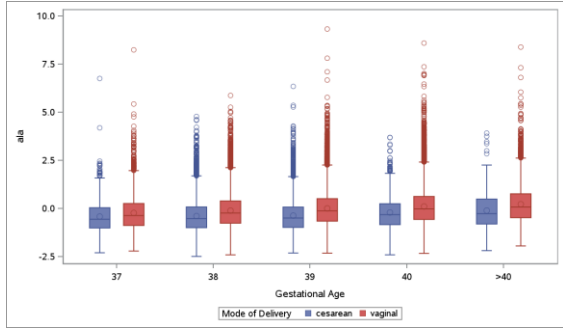
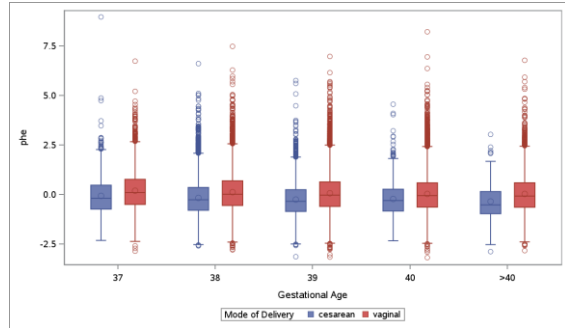
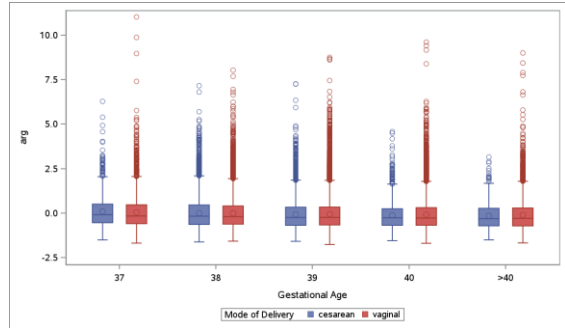
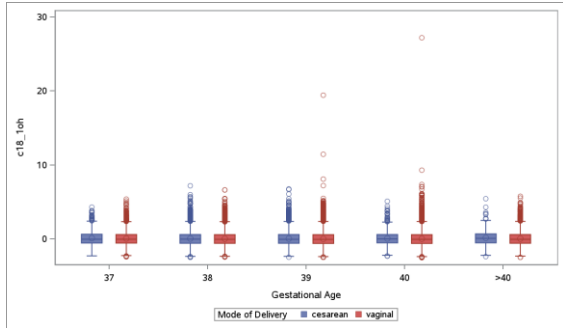
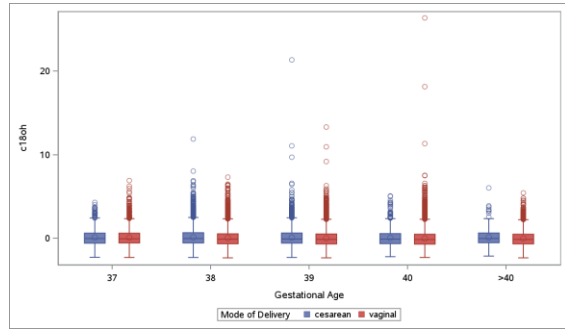
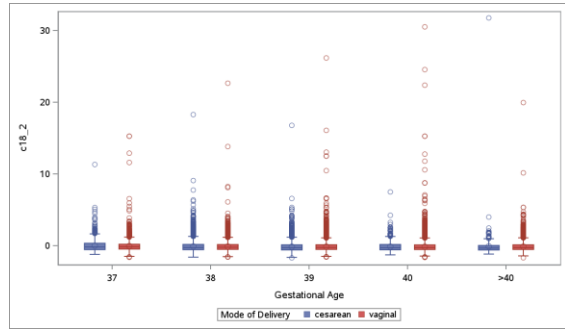
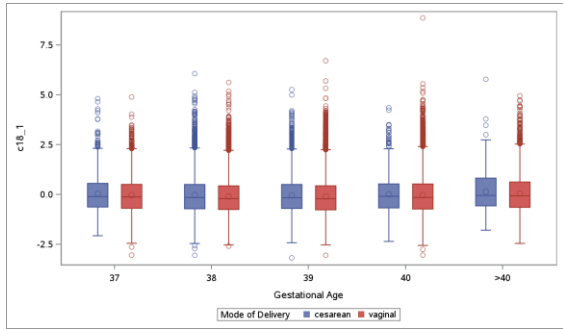


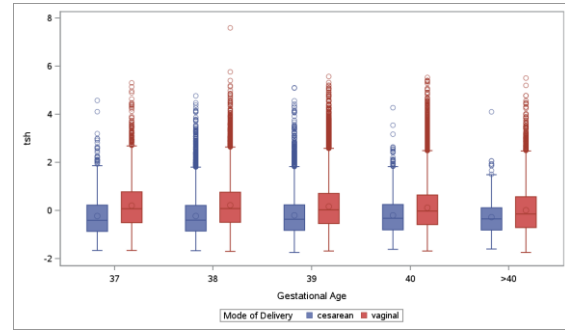
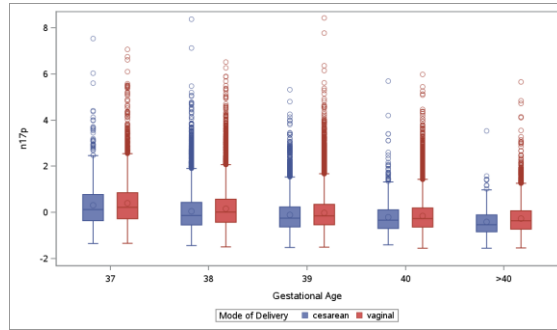
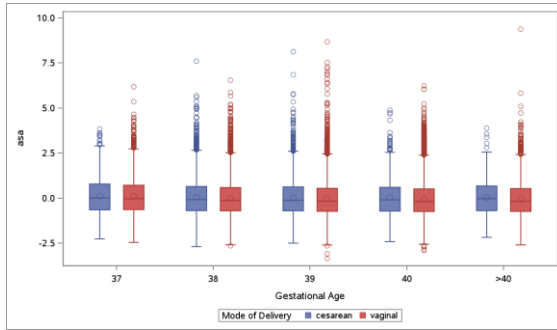
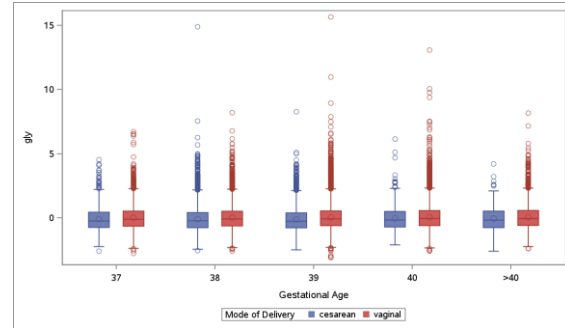
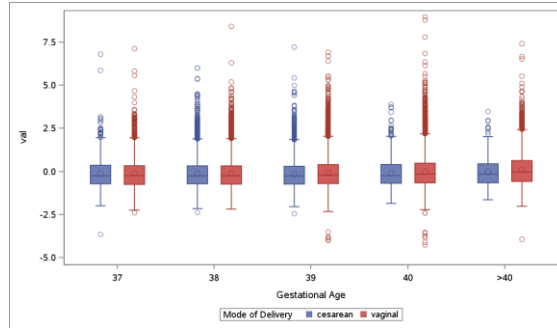
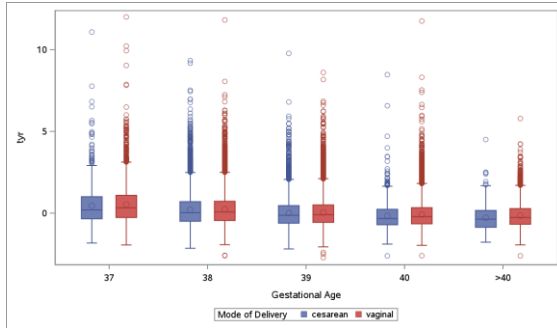
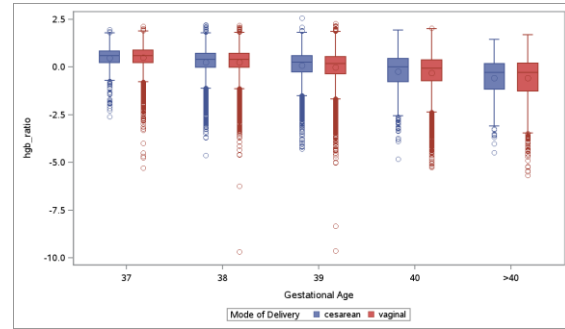
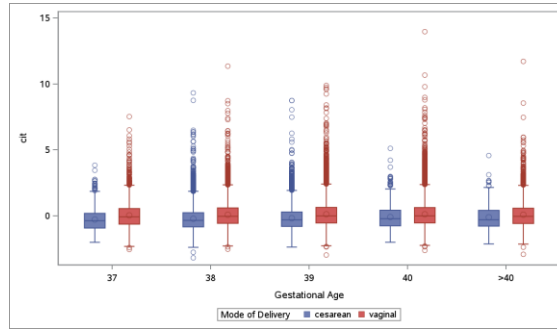
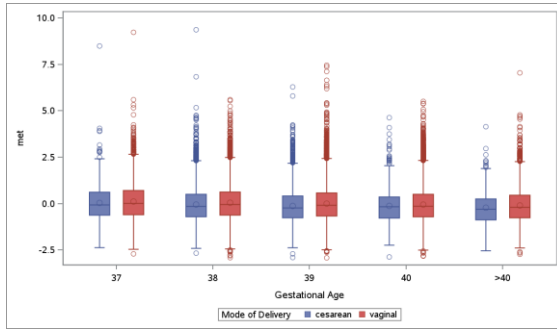
APPENDIX E: Boxplots for Objective 1

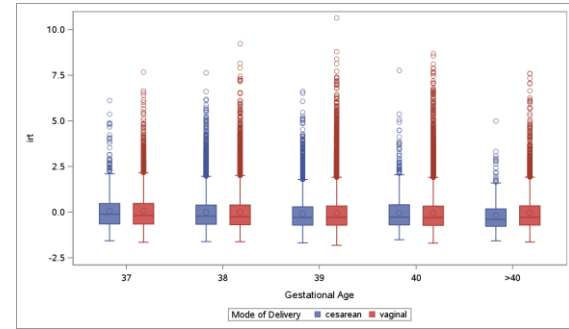
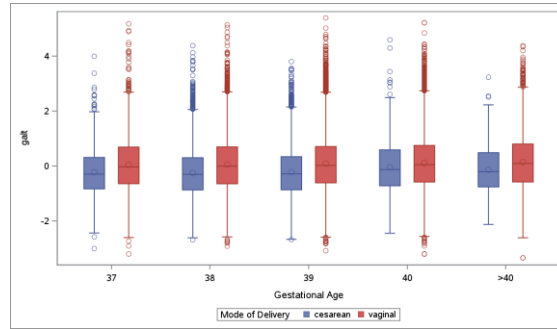
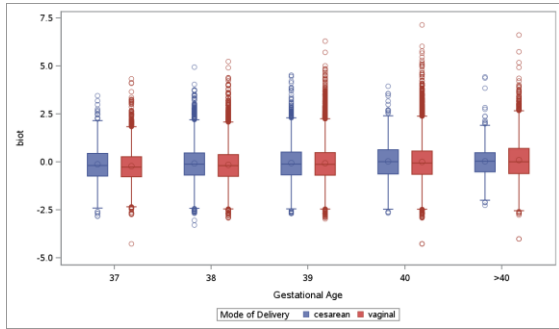




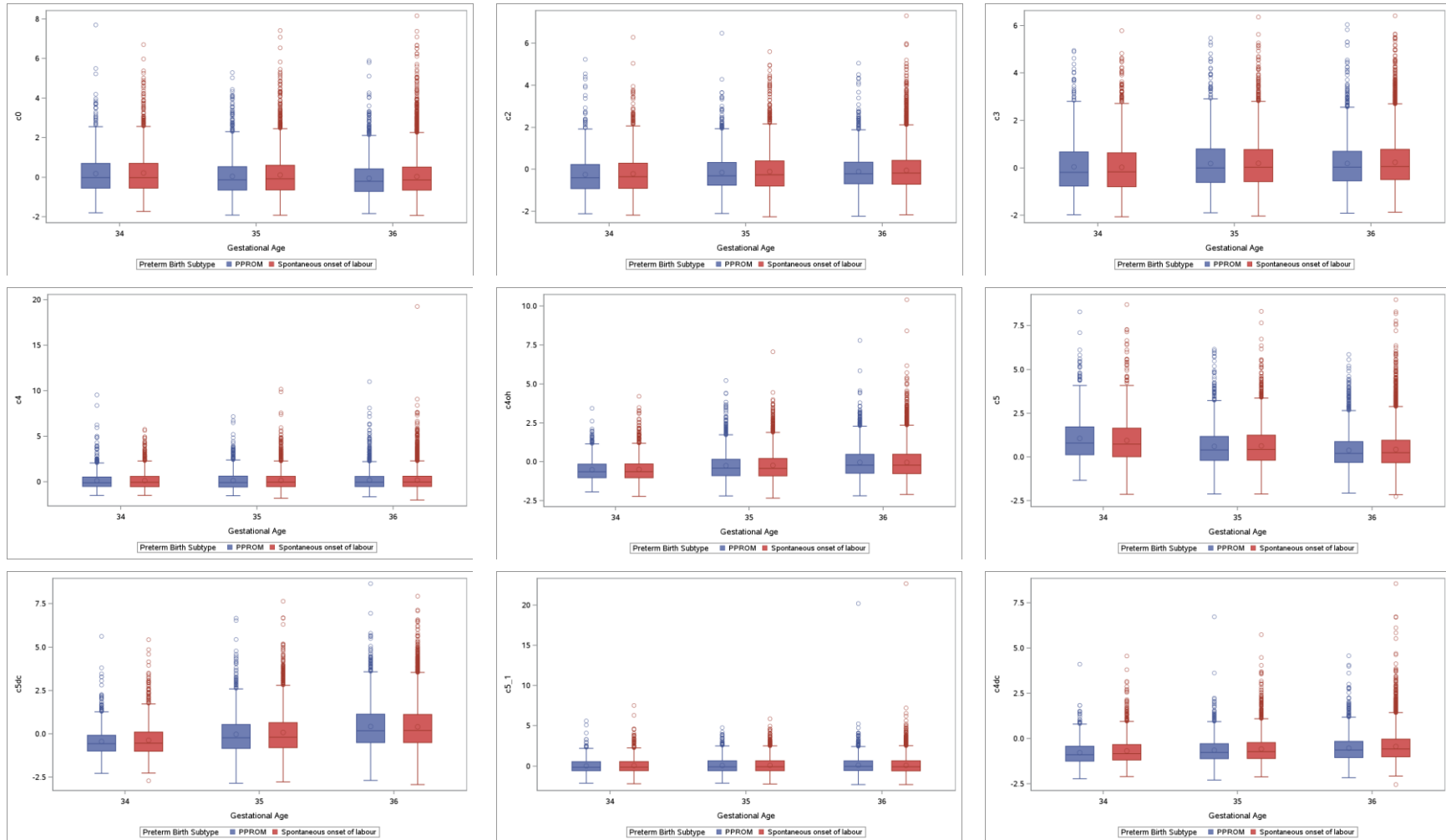


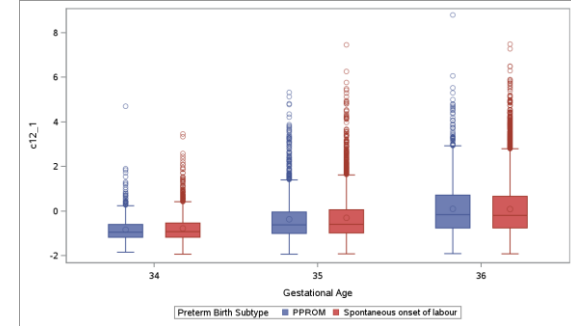
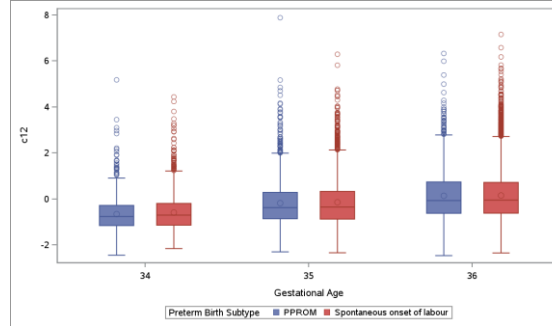
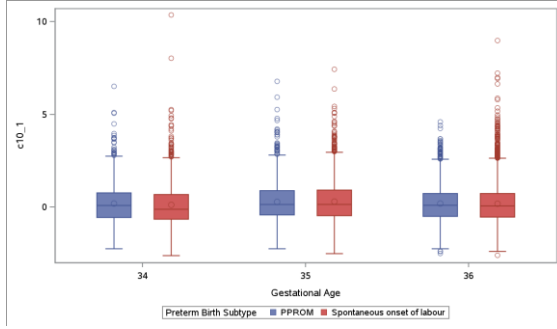
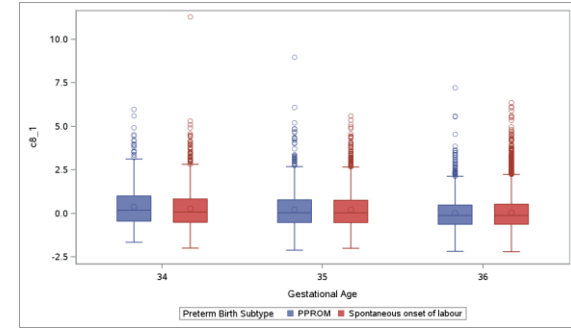
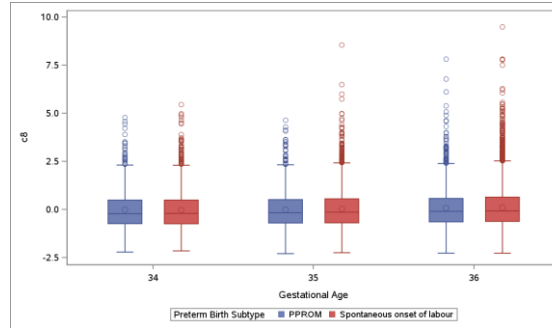
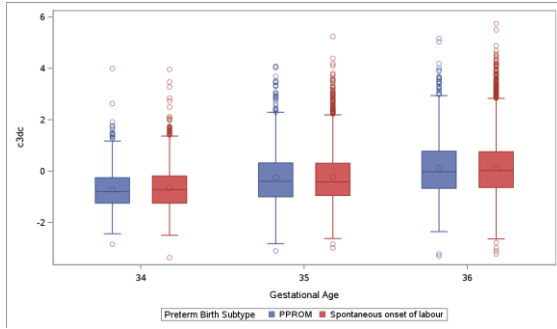
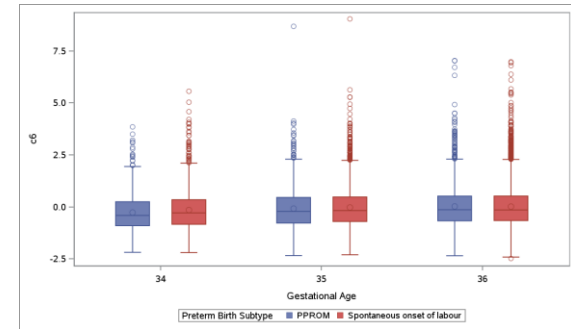
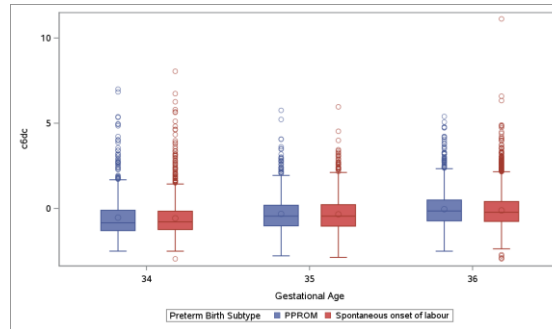
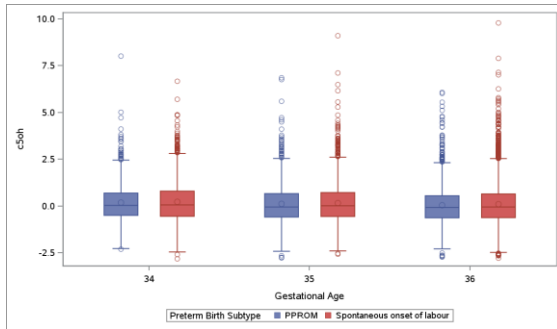


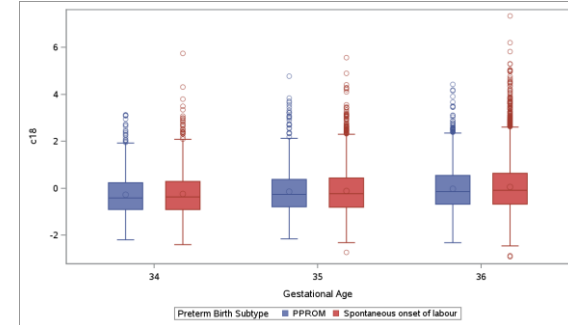
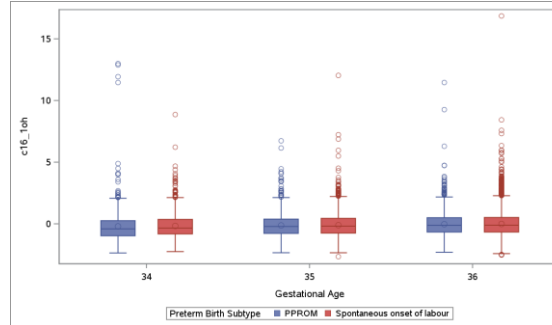
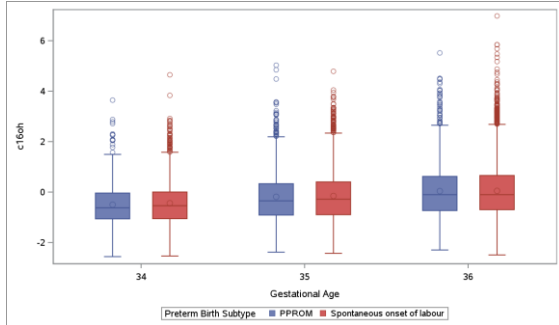
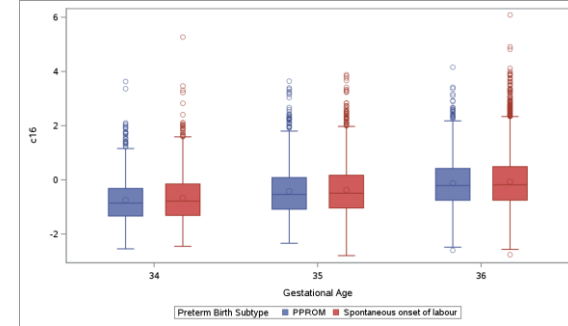
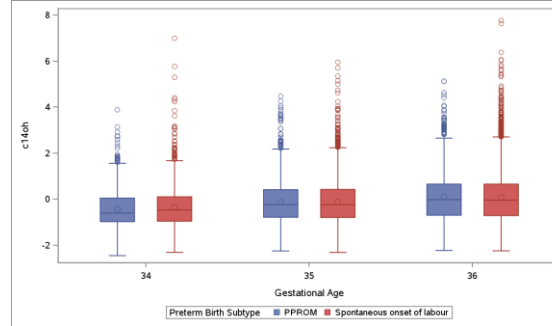
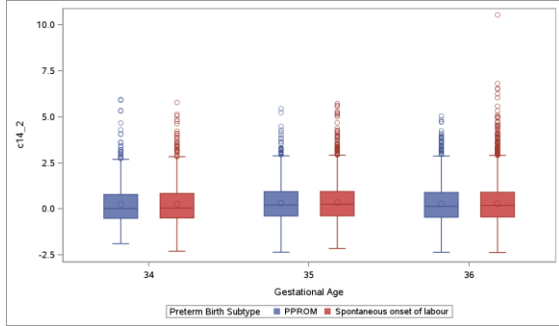
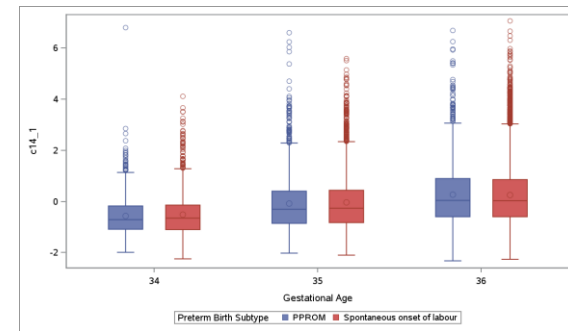
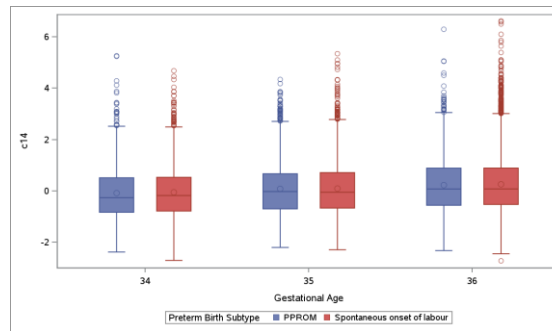
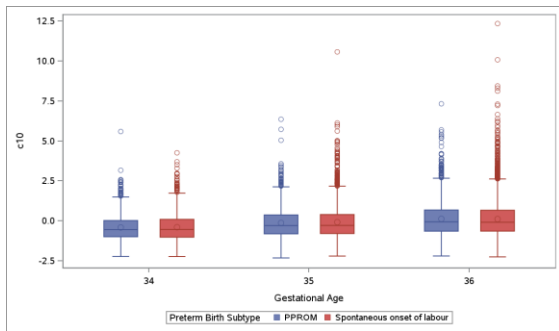


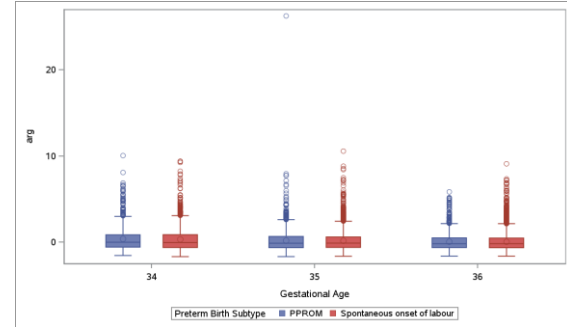
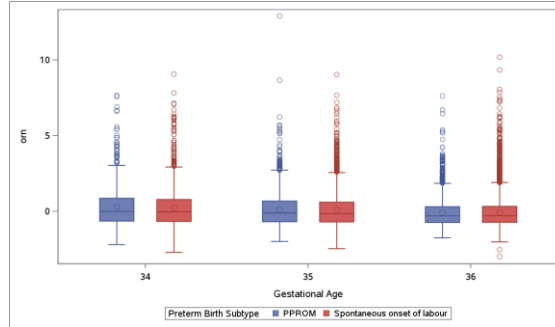
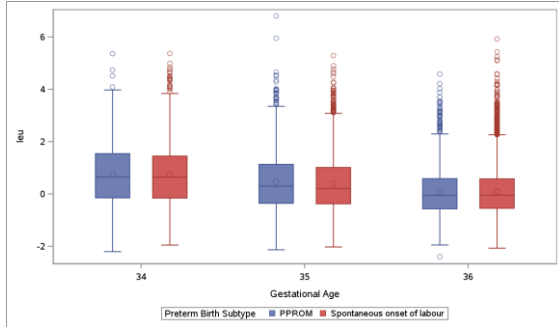
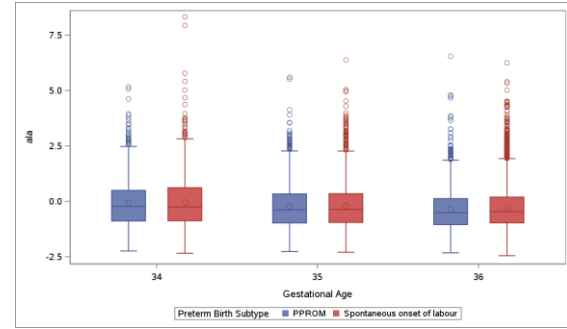
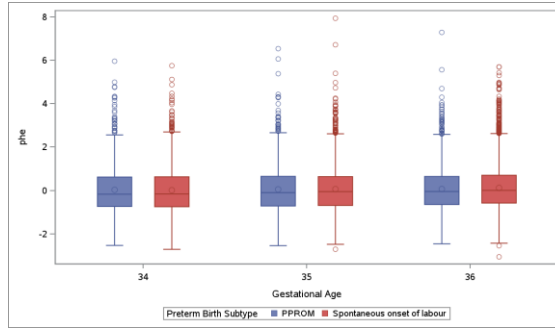
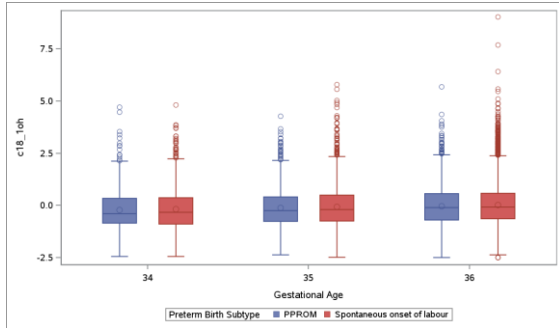
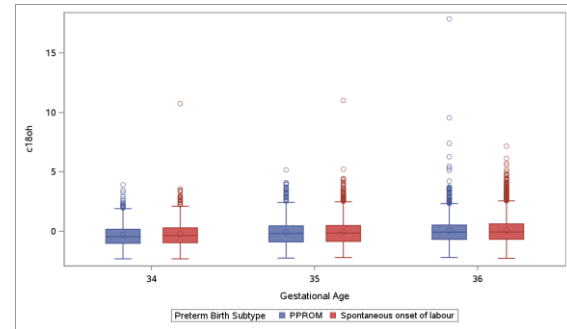
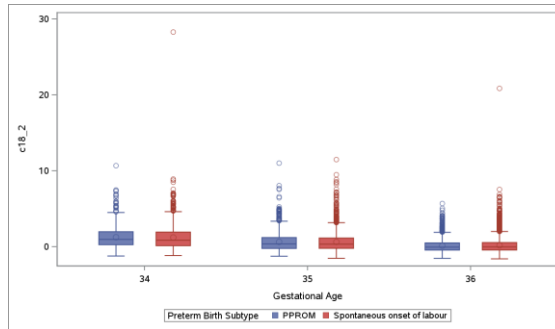
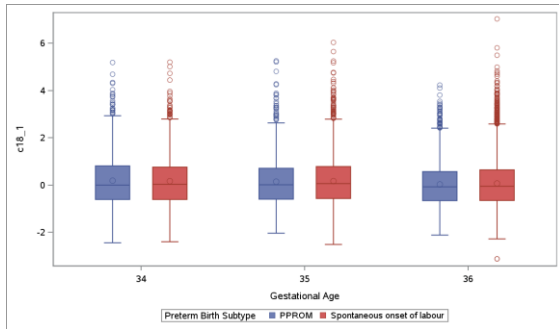


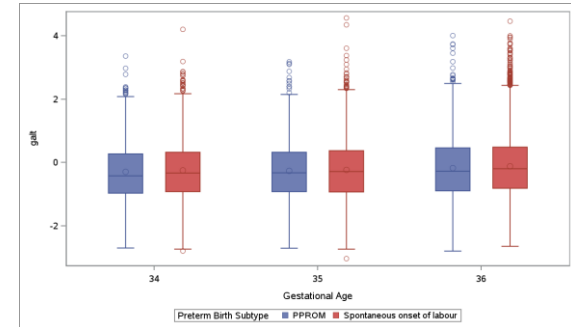
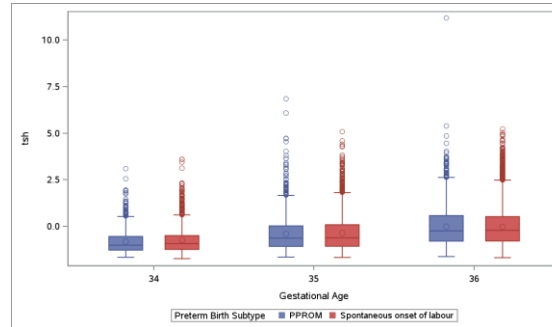
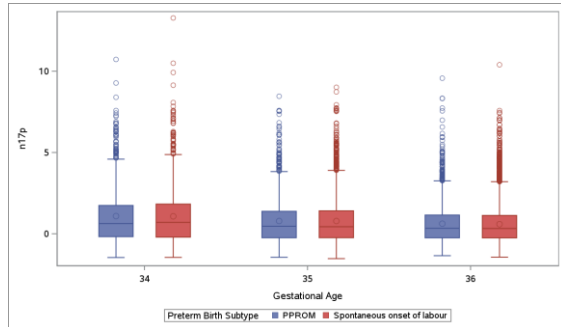
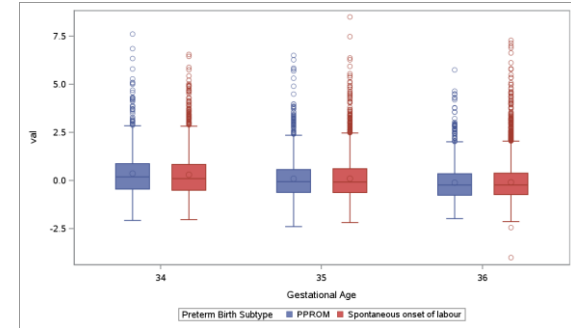
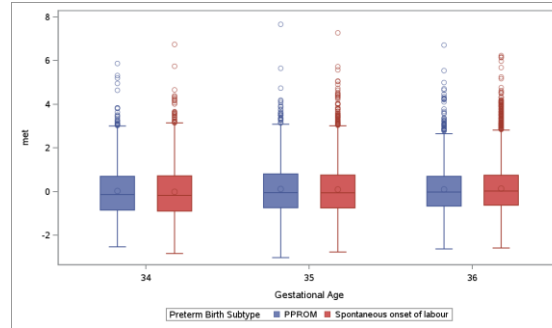
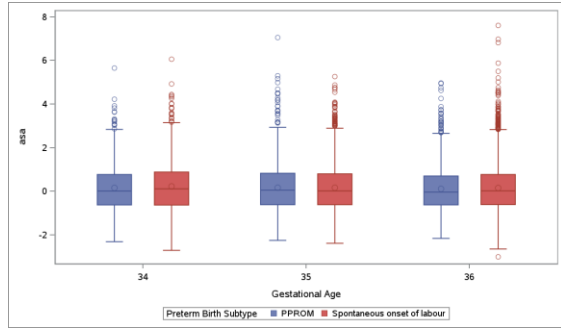
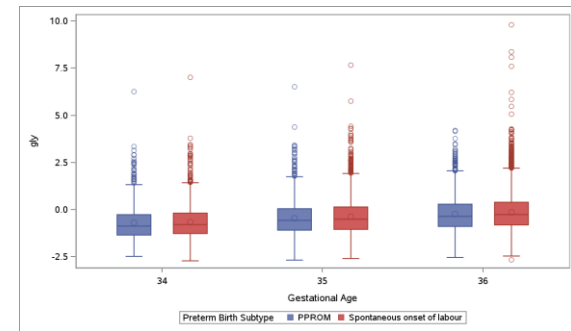
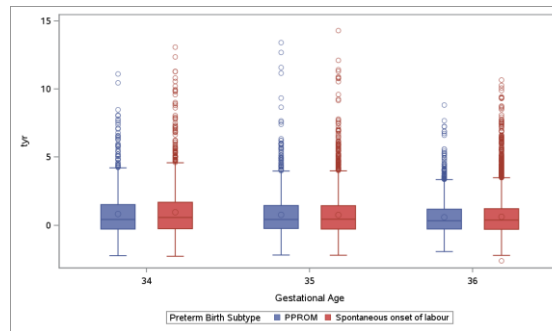
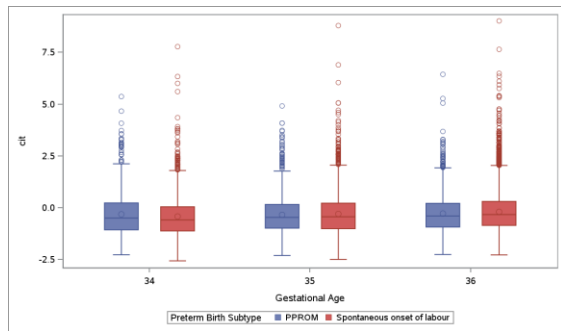
APPENDIX F: Boxplots for Objective 2

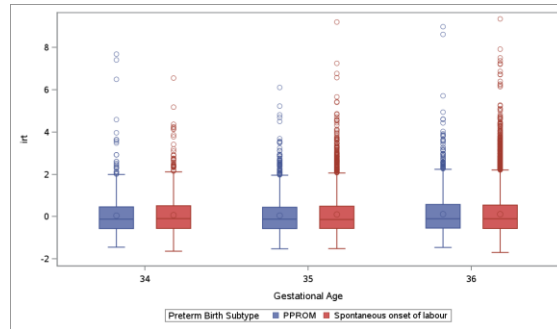
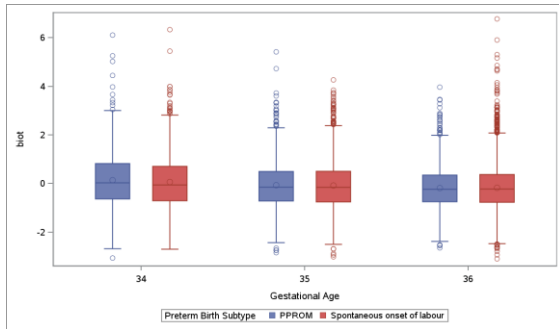












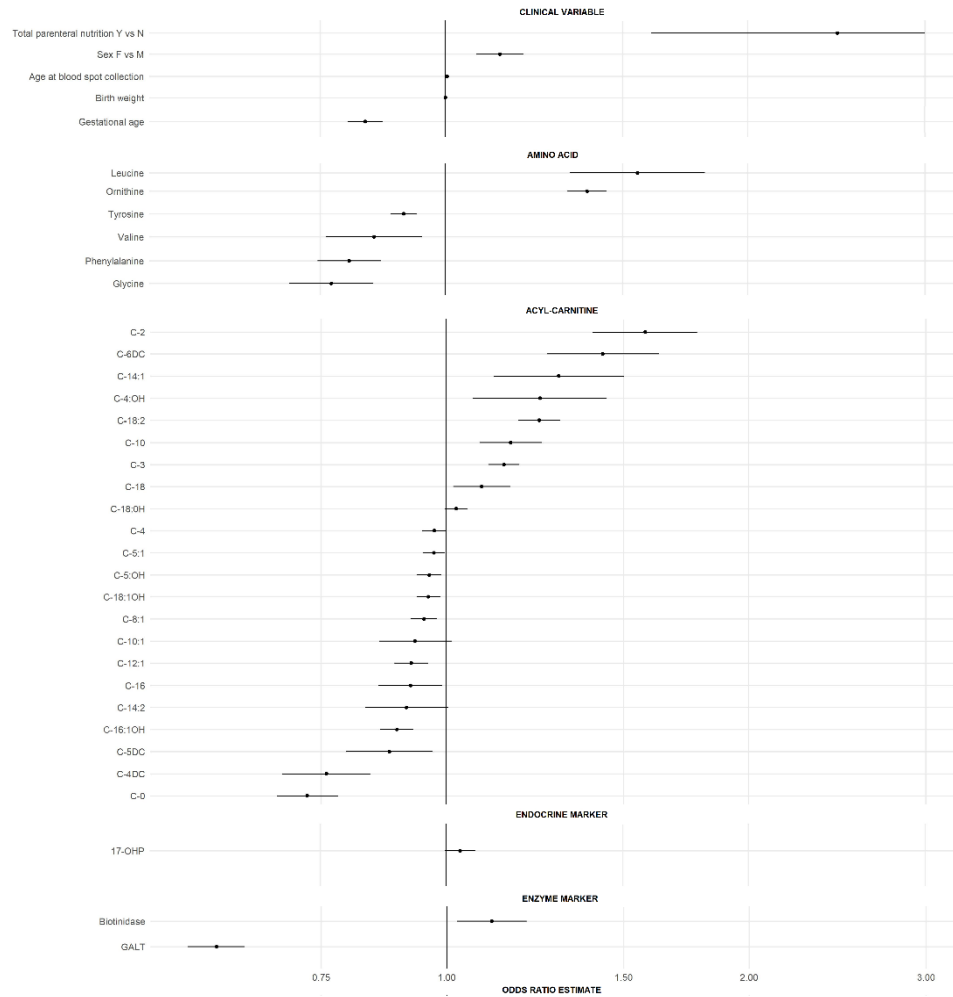
APPENDIX G: Partial Correlation Analysis Results for Objective 1

Rank	Variable	Adjusted rho squared	Rank	Variable	Adjusted rho squared	Rank	Variable	Adjusted rho squared
1	gest	0.084	61	c4dc_c14_2	0.020	121	c14_2	0.004
2	age at collection	0.074	62	c4dc_c5	0.020	122	c10_1	0.003
3	gly	0.050	63	leu_gly	0.020	123	c12	0.002
4	leu_val	0.045	64	c0_c6dc	0.019	124	c5	0.002
5	c3dc_ala	0.042	65	c4dc_c10	0.019	125	c8	0.002
6	c6dc	0.039	66	cit	0.019	126	asa	0.001
7	leu	0.037	67	ala_leu	0.019	127	c12_1	0.001
8	c4oh	0.036	68	ala_biot	0.018	128	c14	0.001
9	c6dc_galt	0.036	69	c10_cit	0.018	129	c14oh	0.001
10	c5dc_gly	0.035	70	c10_galt	0.018	130	c16	0.001
11	c14_gly	0.033	71	c12_galt	0.018	131	c18oh	0.001
12	c5dc_cit	0.032	72	c14_1_c16_1oh	0.018	132	c5oh	0.001
13	c3dc_c4dc	0.031	73	c6dc_cit	0.018	133	met	0.001
14	c4dc_c5dc	0.031	74	c6dc_val	0.018	134	val	0.001
15	ala	0.030	75	c8_ala	0.018	135	arg	0.000
16	c6dc_phe	0.030	76	tsh_irt	0.018	136	biot	0.000
17	c14_1_ala	0.029	77	c12_cit	0.017	137	bweight	0.000
18	c4dc_c14_1	0.029	78	c4dc_c10_1	0.017	138	c18	0.000
19	c5dc_galt	0.029	79	c5dc_c16_1oh	0.017	139	c18_1	0.000
20	leu_cit	0.029	80	c5dc_ala	0.017	140	c18_1oh	0.000
21	phe_leu	0.029	81	c6dc_c16	0.017	141	c18_2	0.000
22	c14_1_cit	0.028	82	c0_c5dc	0.016	142	c2	0.000
23	c14_1_galt	0.028	83	c14_2_cit	0.016	143	c3	0.000
24	c3dc_cit	0.028	84	c2_ala	0.016	144	c4	0.000
25	tsh	0.027	85	c2_c4dc	0.016	145	c5_1	0.000
26	c4oh_ala	0.026	86	c2_c6dc	0.016	146	c6	0.000
27	c5dc_phe	0.026	87	c3_c4dc	0.016	147	c8_1	0.000
28	c3dc_galt	0.025	88	c4dc_c8	0.016	148	irt	0.000

29	c4dc_c4oh	0.025	89	galt	0.016	149	n17p	0.000
30	c4dc_leu	0.025	90	c0_c14_1	0.015	150	orn	0.000
31	c5_ala	0.025	91	c10_1_cit	0.015	151	sex	0.000
32	cit_hgb_ratio	0.025	92	c14_1_gly	0.015	152	tpn	0.000
33	leu_galt	0.025	93	c14_2_galt	0.015	153	tyr	0.000
34	c6dc_gly	0.024	94	c14_galt	0.015			
35	tsh_biot	0.024	95	c3_ala	0.015			
36	c4dc_c18_1	0.023	96	c4oh_phe	0.015			
37	c4dc_hgb_ratio	0.023	97	c5_cit	0.015			
38	c4oh_galt	0.023	98	biot_galt	0.014			
39	ala_met	0.022	99	c14_1_c16	0.014			
40	c10_1_ala	0.022	100	c14_cit	0.014			
41	c18_1_ala	0.022	101	c14oh_ala	0.014			
42	c3dc_phe	0.022	102	c18_1_galt	0.014			
43	c6dc_c16_1oh	0.022	103	c3dc_c16_1oh	0.014			
44	ala_gly	0.021	104	c3dc_gly	0.014			
45	c10_ala	0.021	105	c4dc_biot	0.014			
46	c12_ala	0.021	106	c4dc_c18	0.014			
47	c16_1oh	0.021	107	c4oh_c16_1oh	0.014			
48	c16oh	0.021	108	c5_galt	0.014			
49	c16oh_ala	0.021	109	c6dc_met	0.014			
50	c4dc_c12	0.021	110	c8_cit	0.014			
51	c4oh_cit	0.021	111	orn_cit	0.014			
52	hgb_ratio_galt	0.021	112	phe	0.013			
53	phe_hgb_ratio	0.021	113	hgb_ratio	0.011			
54	ala_orn	0.020	114	c14_1	0.009			
55	ala_val	0.020	115	c5dc	0.009			
56	c14_1_phe	0.020	116	c6dc_ala	0.007			
57	c14_2_ala	0.020	117	c3dc	0.006			
58	c14_ala	0.020	118	ala_hgb_ratio	0.006			
59	c4dc	0.020	119	c0	0.004			
60	c4dc_c14	0.020	120	c10	0.004			

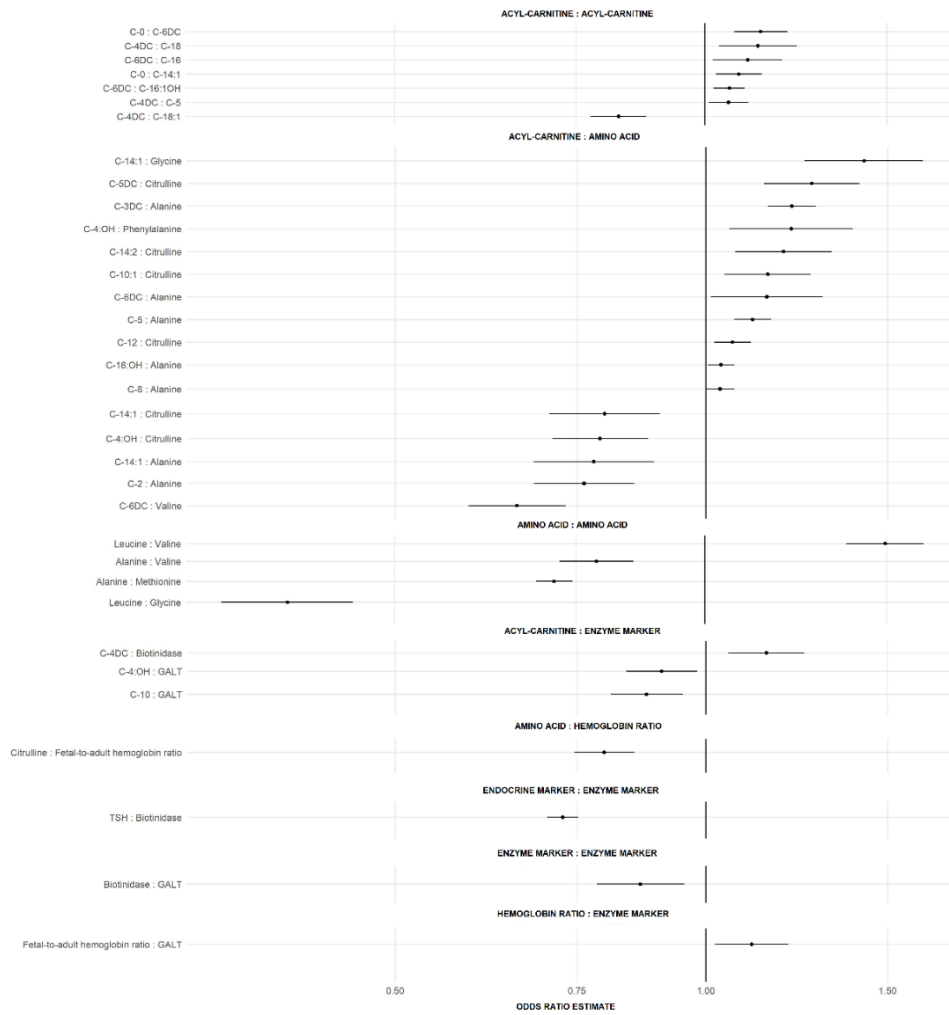
APPENDIX H: Forest Plots for Sensitivity Analyses for Objective 1

Supplemental Figure H.1a Individual analytes and clinical variables in the final model of early-term infants



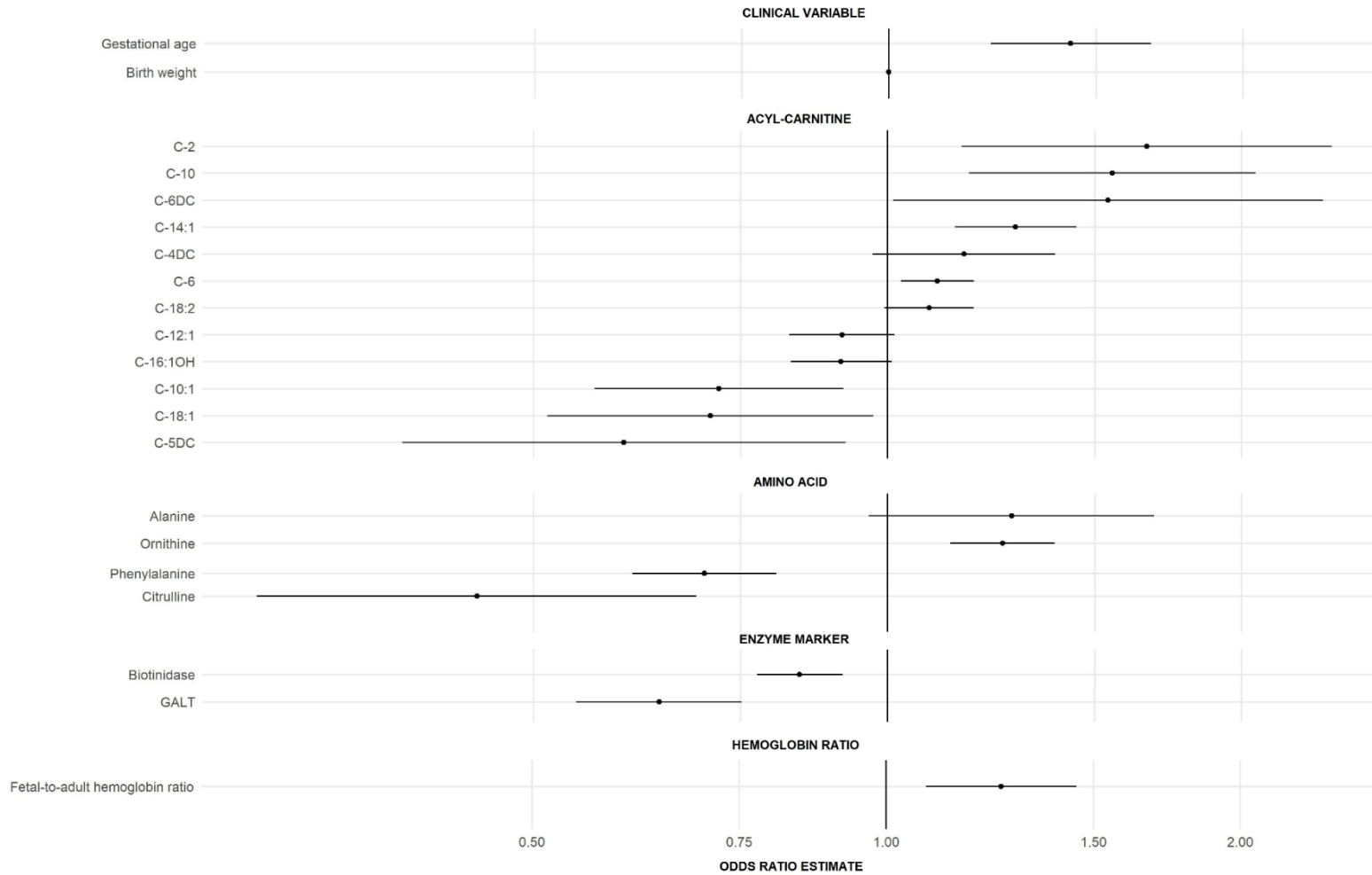
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.1b Analyte ratios in the final model of early-term infants



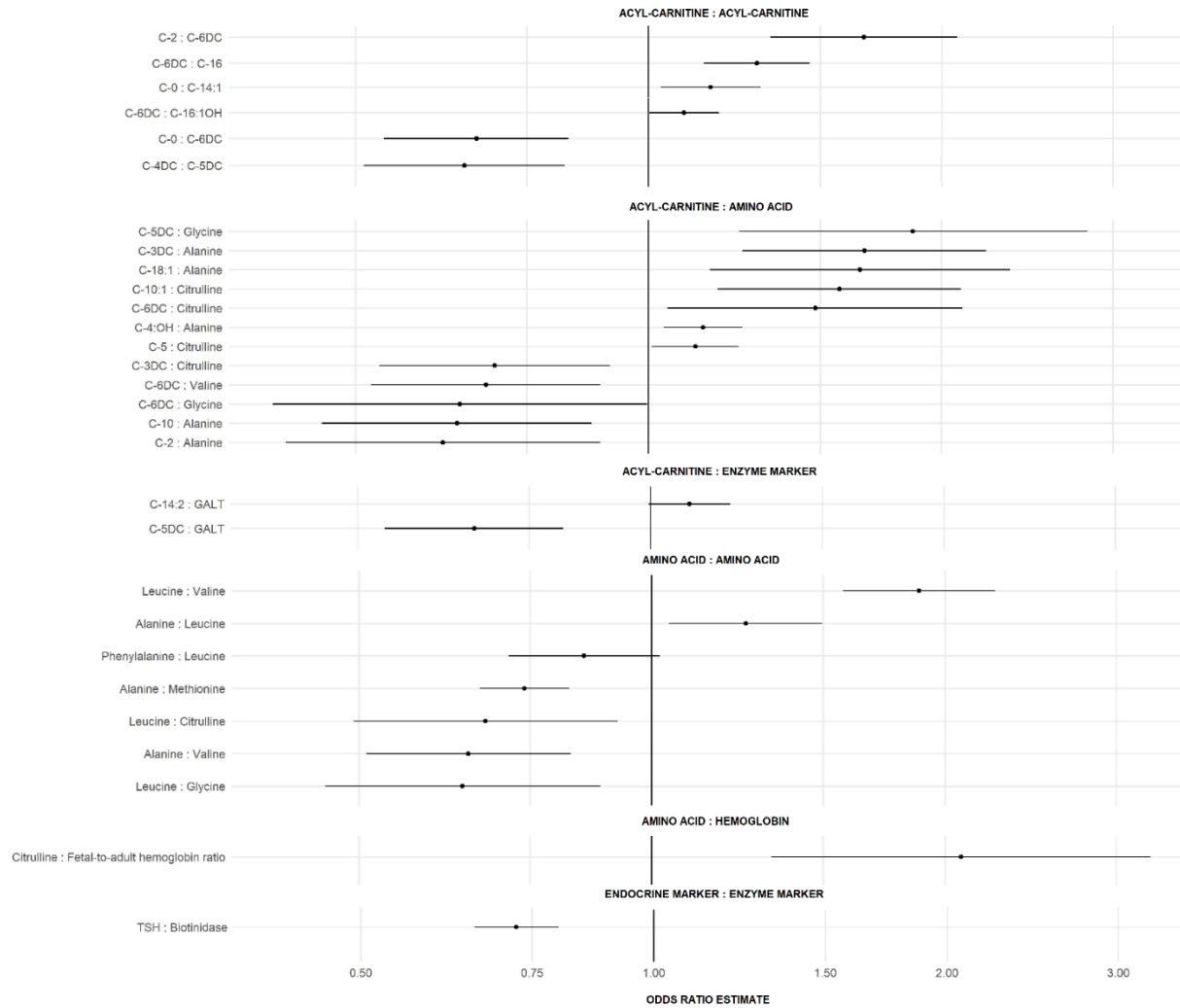
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.2a Individual analytes and clinical variables in the final model of full-term infants



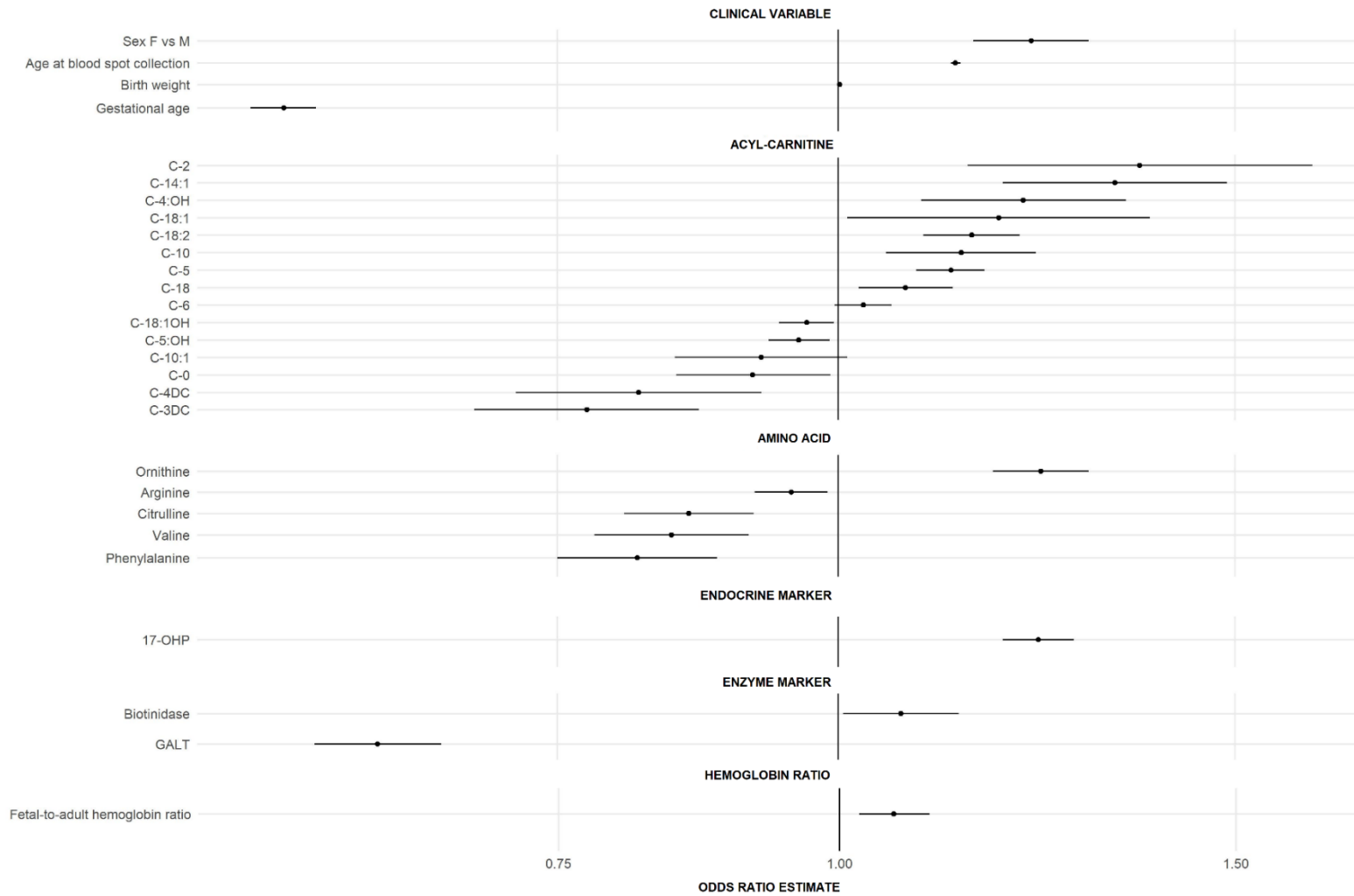
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.2b Analyte ratios in the final model of full-term infants



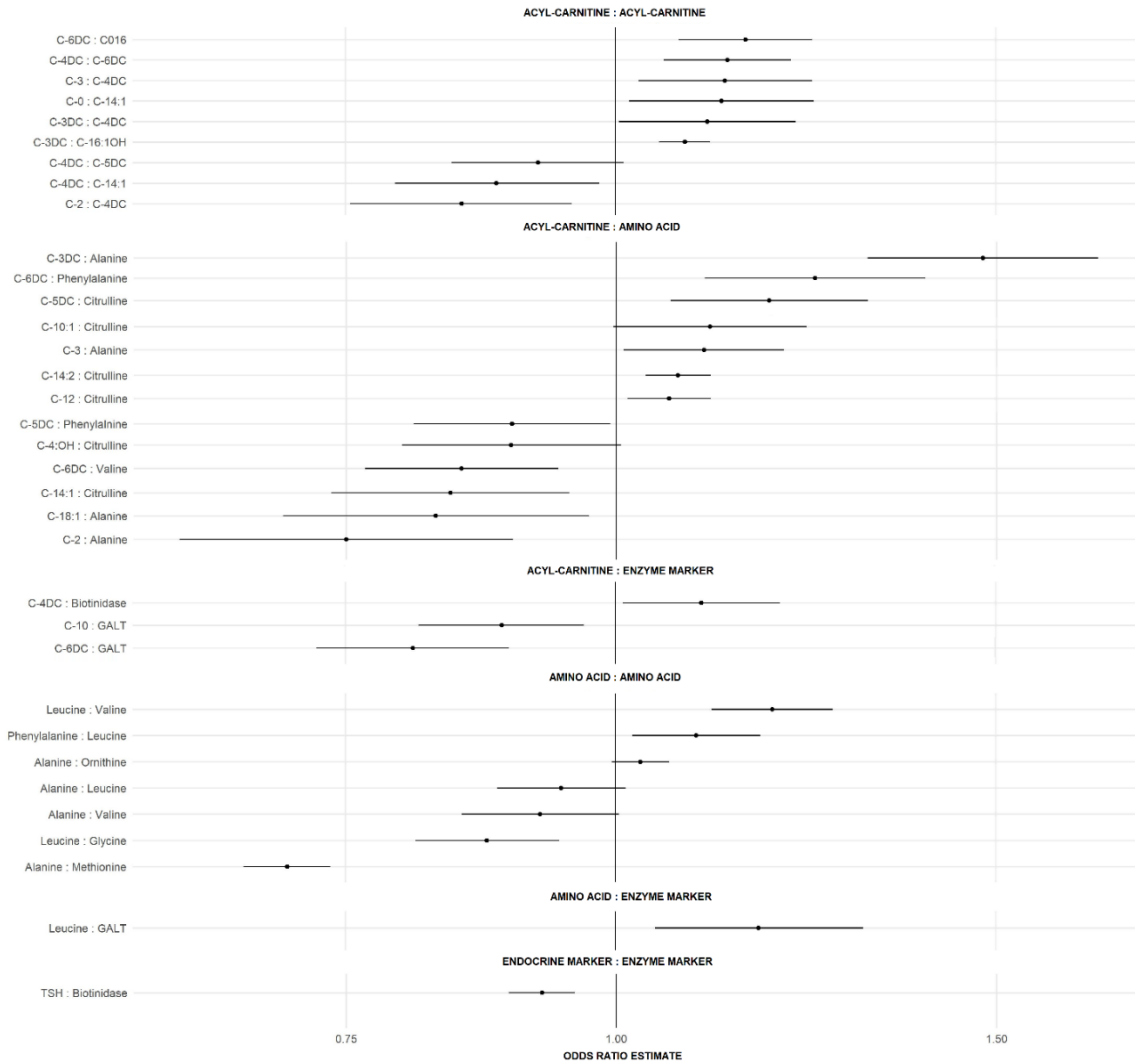
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.3a Individual analytes and clinical variables in the final model of-term infants born within recommended window



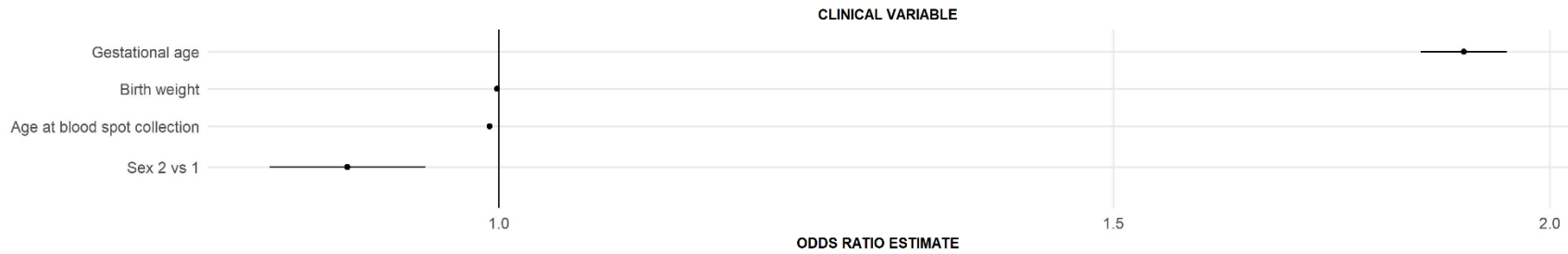
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.3b Analyte ratios in the final model of term infants born within recommended window



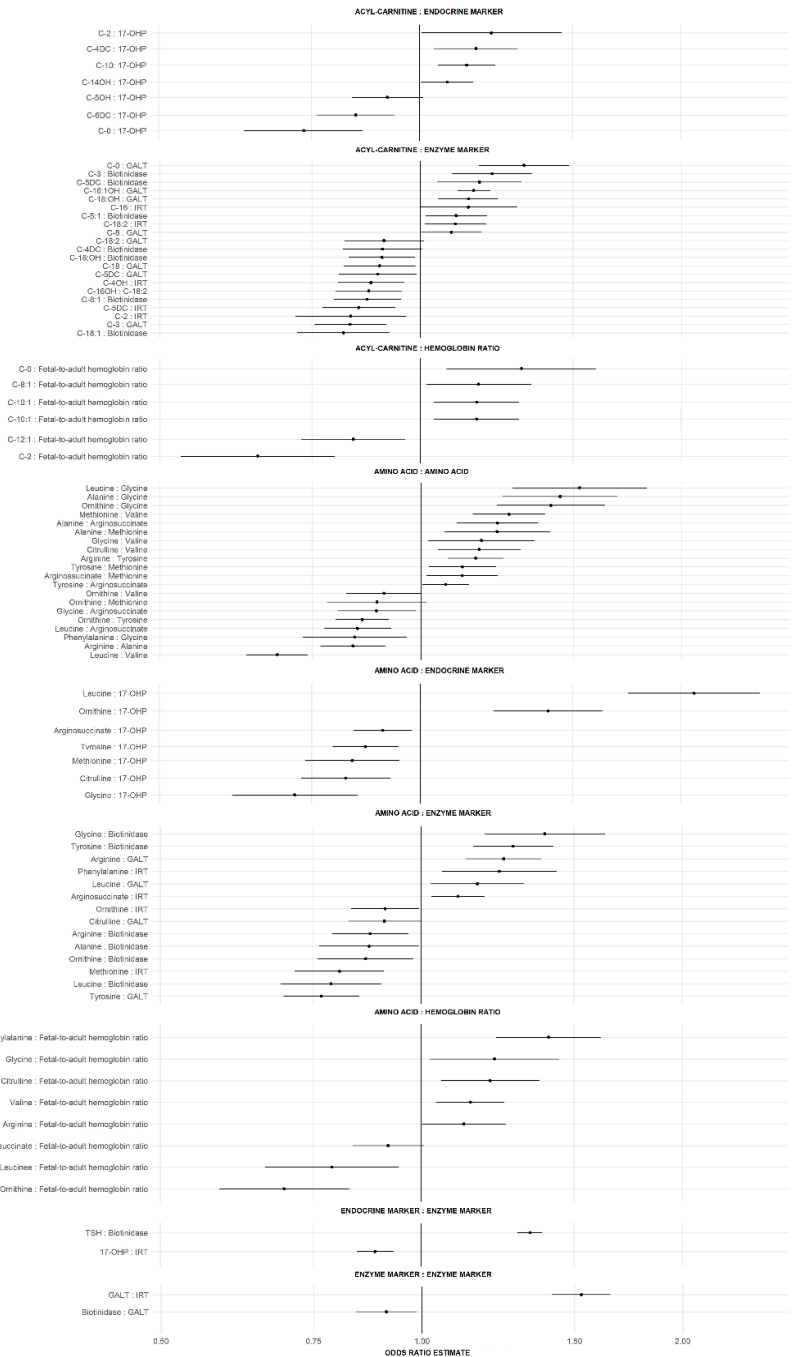
Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.4a Individual analytes and clinical variables in the final model of-term infants with the inclusion of remaining ranked analyte ratios based on Spearman's correlation until the maximum number of parameters was reached



Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

Supplemental Figure H.4b Analyte ratios in the final model of term infants with the inclusion of remaining ranked analyte ratios based on Spearman's correlation until the maximum number of parameters was reached



Forest plot demonstrating associations of newborn screening analytes with planned cesarean delivery with no labour, compared with vaginal delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with planned cesarean delivery, while ORs less than one indicate an inverse association with planned cesarean delivery.

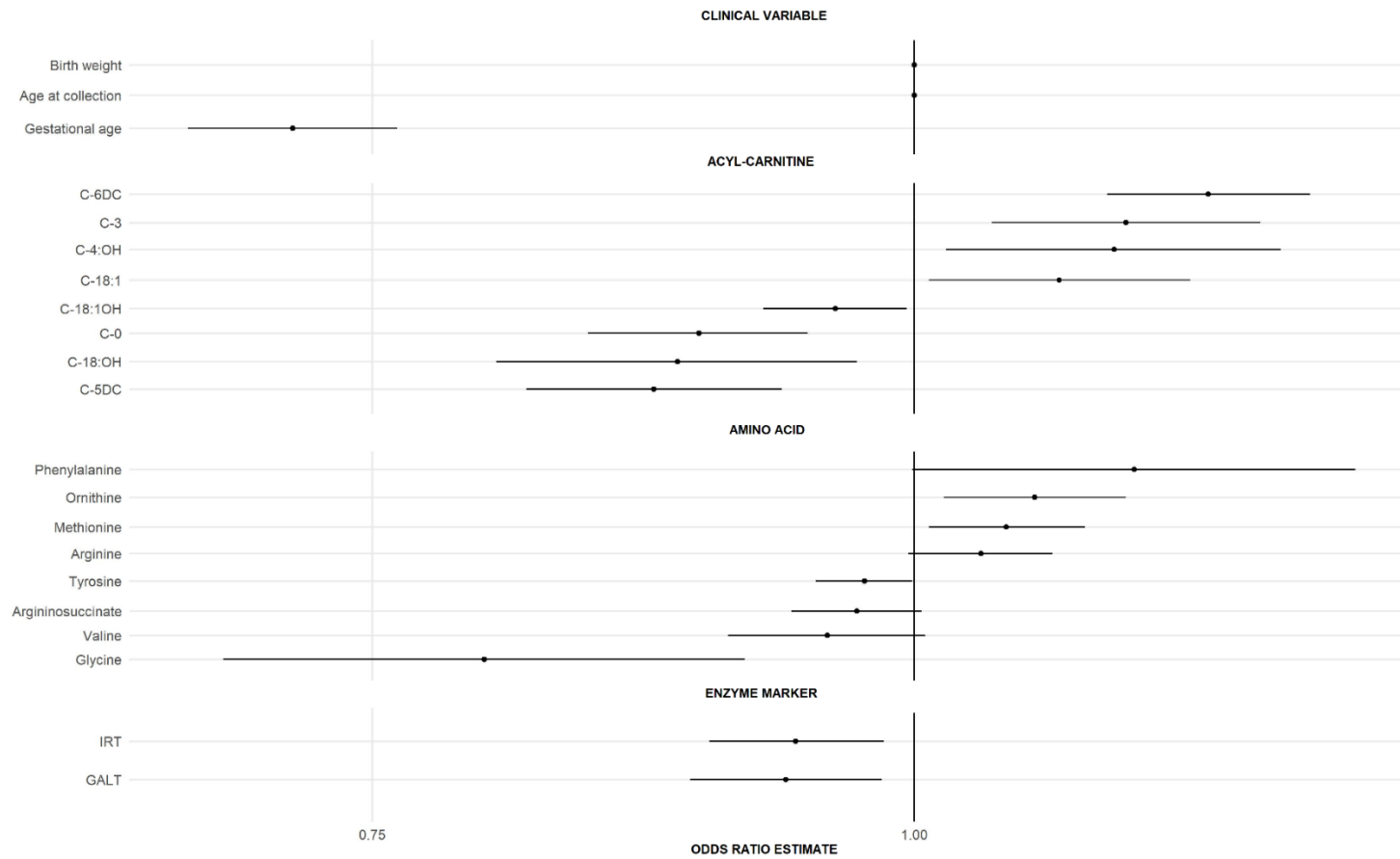
APPENDIX I: Partial Correlation Analysis for Objective 2

Rank	Variable	Adjusted rho squared	Rank	Variable	Adjusted rho squared	Rank	Variable	Adjusted rho squared
1	gest	0.02	61	c12_1_phe	0.002	121	c18oh	0.001
2	ageatcollection	0.01	62	c18_2_galt	0.002	122	c18_1oh	0.001
3	bweight	0.006	63	c18_2_cit	0.002	123	leu	0.001
4	c5dc_c18_2	0.006	64	c14_1_c18_2	0.002	124	cit	0.001
5	c2_c18_2	0.005	65	c3dc_val	0.002	125	gly	0.001
6	c12_1_val	0.004	66	c12_biot	0.002	126	galt	0.001
7	c4dc_leu	0.004	67	c18_leu	0.002	127	sex	0
8	gly_val	0.004	68	c14_c18_2	0.002	128	tpn	0
9	c3dc_biot	0.004	69	c18_2_met	0.002	129	delttype	0
10	c18oh_leu	0.004	70	phe_gly	0.002	130	c0	0
11	c4dc	0.003	71	c5dc_val	0.002	131	c3	0
12	tsh_biot	0.003	72	c5dc_biot	0.002	132	c4	0
13	c16	0.003	73	c18_2_c18_1oh	0.002	133	c5	0
14	tsh	0.003	74	ala_gly	0.002	134	c5_1	0
15	c12_1_leu	0.003	75	c18_2_ala	0.002	135	c5oh	0
16	c4dc_c5	0.003	76	c5_c12_1	0.002	136	c6dc	0
17	c16_c18_2	0.003	77	c18_1_c18_2	0.002	137	c8	0
18	c4dc_biot	0.003	78	c12_val	0.002	138	c8_1	0
19	c4dc_c8_1	0.003	79	c5oh_c18_2	0.002	139	c10_1	0
20	c18_c18_2	0.003	80	c6_c18_2	0.002	140	c14_2	0
21	c0_c18_2	0.003	81	c4dc_ala	0.002	141	c14oh	0
22	c3dc_leu	0.003	82	c18_2_phe	0.002	142	c18_1	0
23	c4dc_val	0.003	83	c16oh_phe	0.002	143	arg	0
24	c16_c18_1	0.003	84	c3_c18_2	0.002	144	phe	0
25	c5_c12_1	0.003	85	c12_1_met	0.002	145	ala	0
26	c10_1_c12_1	0.003	86	c5_c5dc	0.002	146	orn	0
27	orn_gly	0.003	87	c4dc_phe	0.002	147	tyr	0
28	c18_2_c18oh	0.003	88	c3dc_c5	0.002	148	asa	0
29	c4dc_c10_1	0.003	89	c8_c12_1	0.002	149	met	0

30	c12_leu	0.003	90	c14_1_val	0.002	150	val	0
31	c4dc_c18_2	0.003	91	c8_1_c18oh	0.002	151	n17p	0
32	c12_1	0.003	92	c8_1_gly	0.002	152	biot	0
33	c12_c18_2	0.003	93	c4dc_orn	0.002	153	irt	0
34	c12_1_biot	0.003	94	c5oh_c12_1	0.002			
35	c12_1_c18_2	0.003	95	c5_c14_1	0.002			
36	c5dc_leu	0.003	96	c18oh_leu	0.002			
37	c3dc_c18_2	0.003	97	c12_1_c14	0.002			
38	c14_1_leu	0.003	98	c12_1_ala	0.002			
39	c5_c16	0.003	99	c16_1oh_c18_2	0.002			
40	c16oh_c18_2	0.003	100	c14_1_biot	0.002			
41	leu_cit	0.003	101	c8_1_c16	0.002			
42	c8_1_c12_1	0.003	102	c5dc_c12_1	0.002			
43	c16_biot	0.003	103	c4oh_c18_2	0.002			
44	leu_gly	0.003	104	c10_1_gly	0.002			
45	gly_biot	0.003	105	c10_c18_2	0.002			
46	c6dc_c12_1	0.003	106	c12_1_orn	0.002			
47	c12_1_n17p	0.003	107	c4dc_met	0.002			
48	c16_val	0.003	108	tsh_galt	0.002			
49	c16_leu	0.003	109	c10_1_c18oh	0.002			
50	phe_leu	0.003	110	c12_1_arg	0.002			
51	c3dc	0.002	111	c4_c12_1	0.002			
52	c5dc	0.002	112	c3dc_c8_1	0.002			
53	c12	0.002	113	c2_c18_2	0.001			
54	c14_1	0.002	114	c4oh	0.001			
55	c18_2	0.002	115	c6	0.001			
56	c5_gly	0.002	116	c10	0.001			
57	leu_galt	0.002	117	c14	0.001			
58	tsh_irt	0.002	118	c16oh	0.001			
59	c4dc_c18_1	0.002	119	c16_1oh	0.001			
60	c12_1_c18_1	0.002	120	c18	0.001			

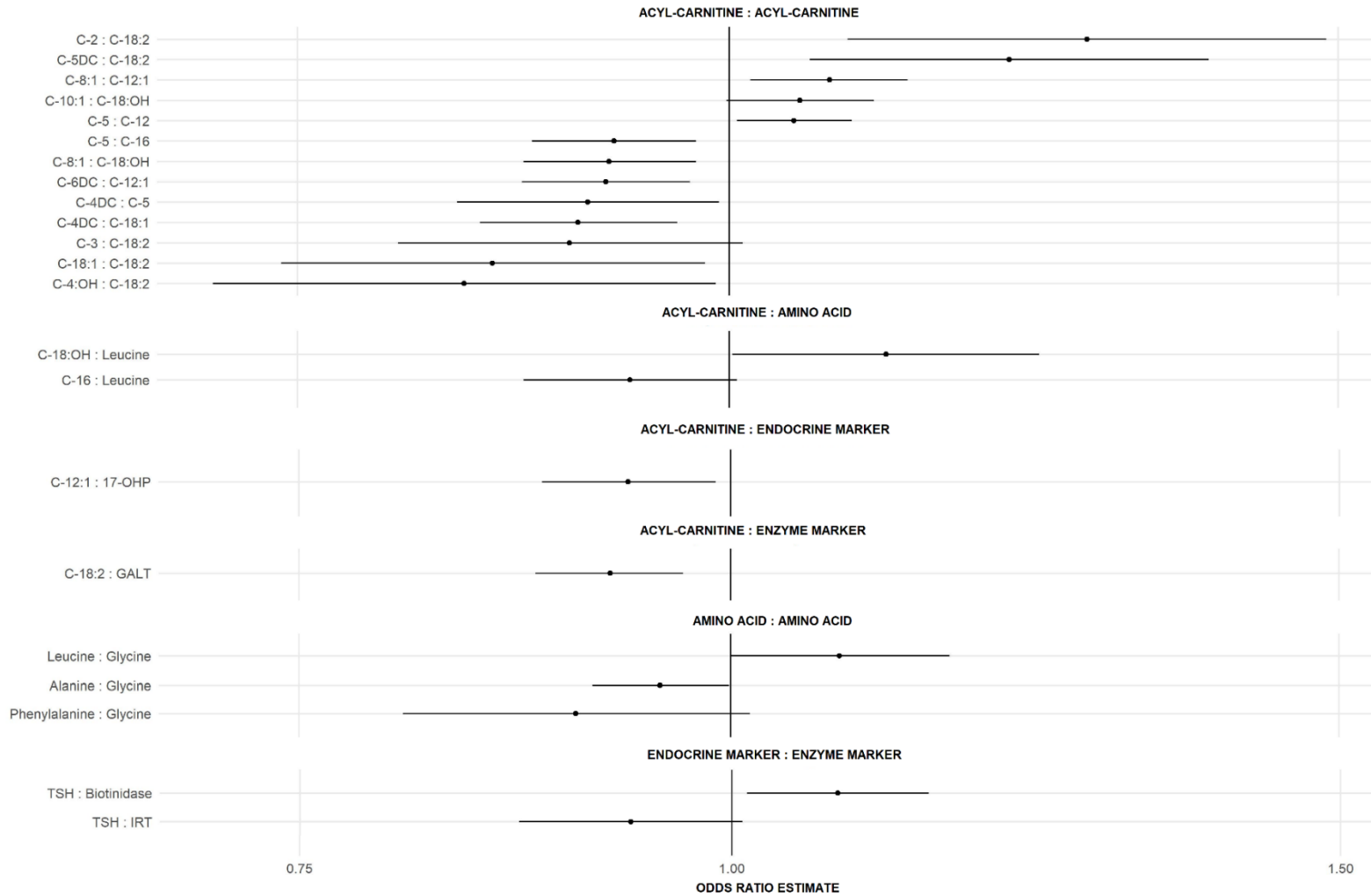
APPENDIX J: Forest Plots for Sensitivity Analyses for Objective 2

Supplemental Figure J.1a Individual analytes and clinical variables in the final model of preterm, including those with transfusions



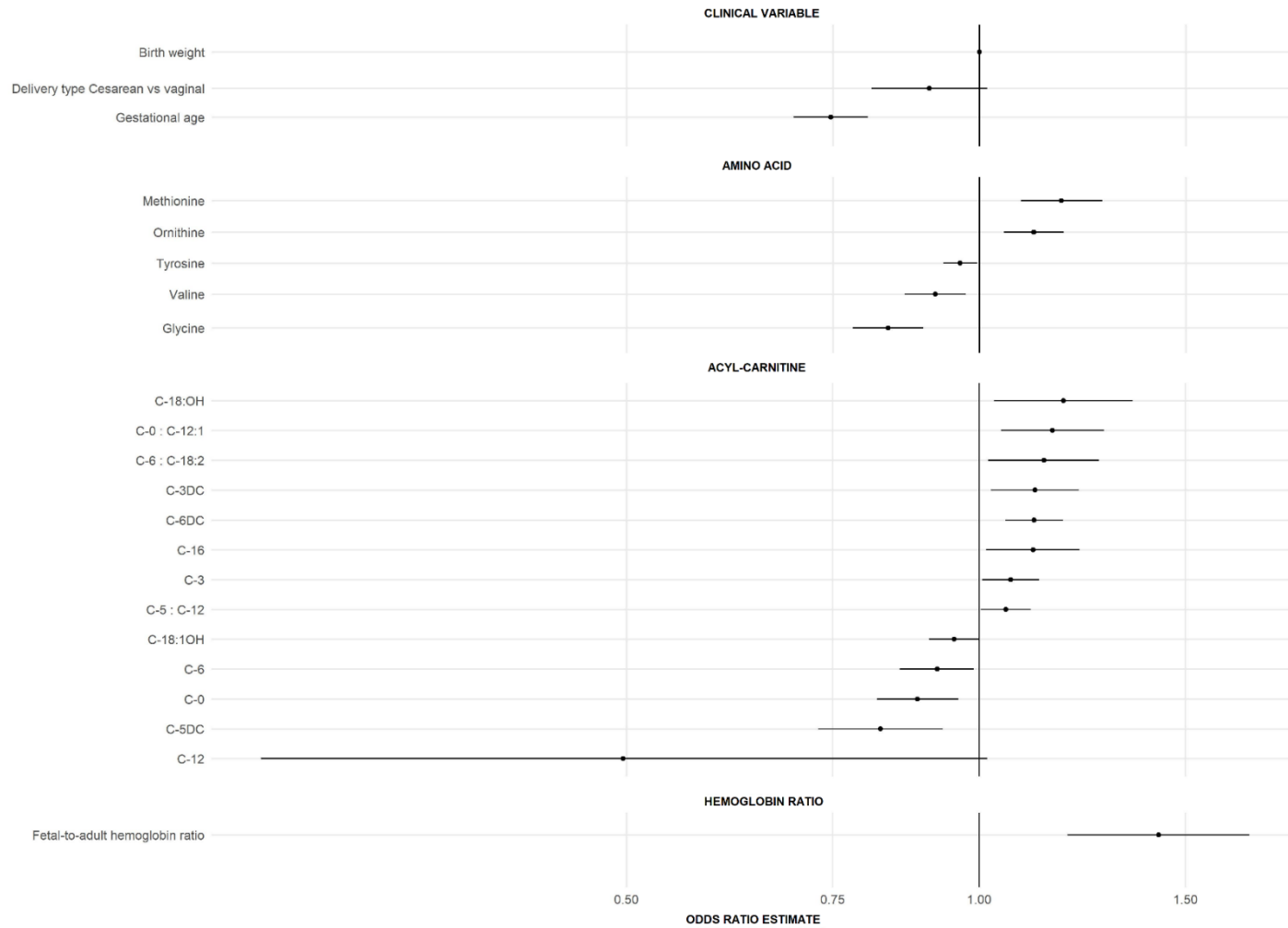
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.1b Analyte ratios in the final model of preterm infants, including those with transfusions



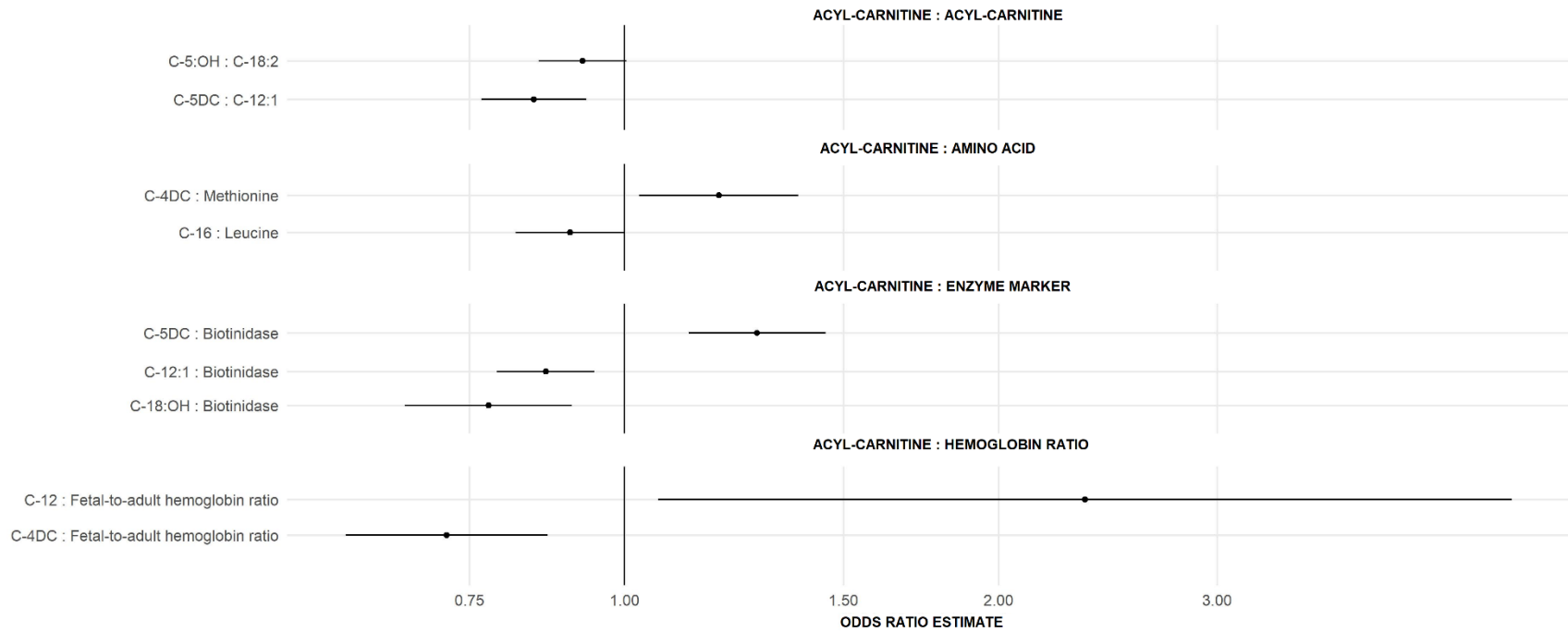
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.2a Individual analytes and clinical variables in the final model of pre-term infants born between 2010 and 2012



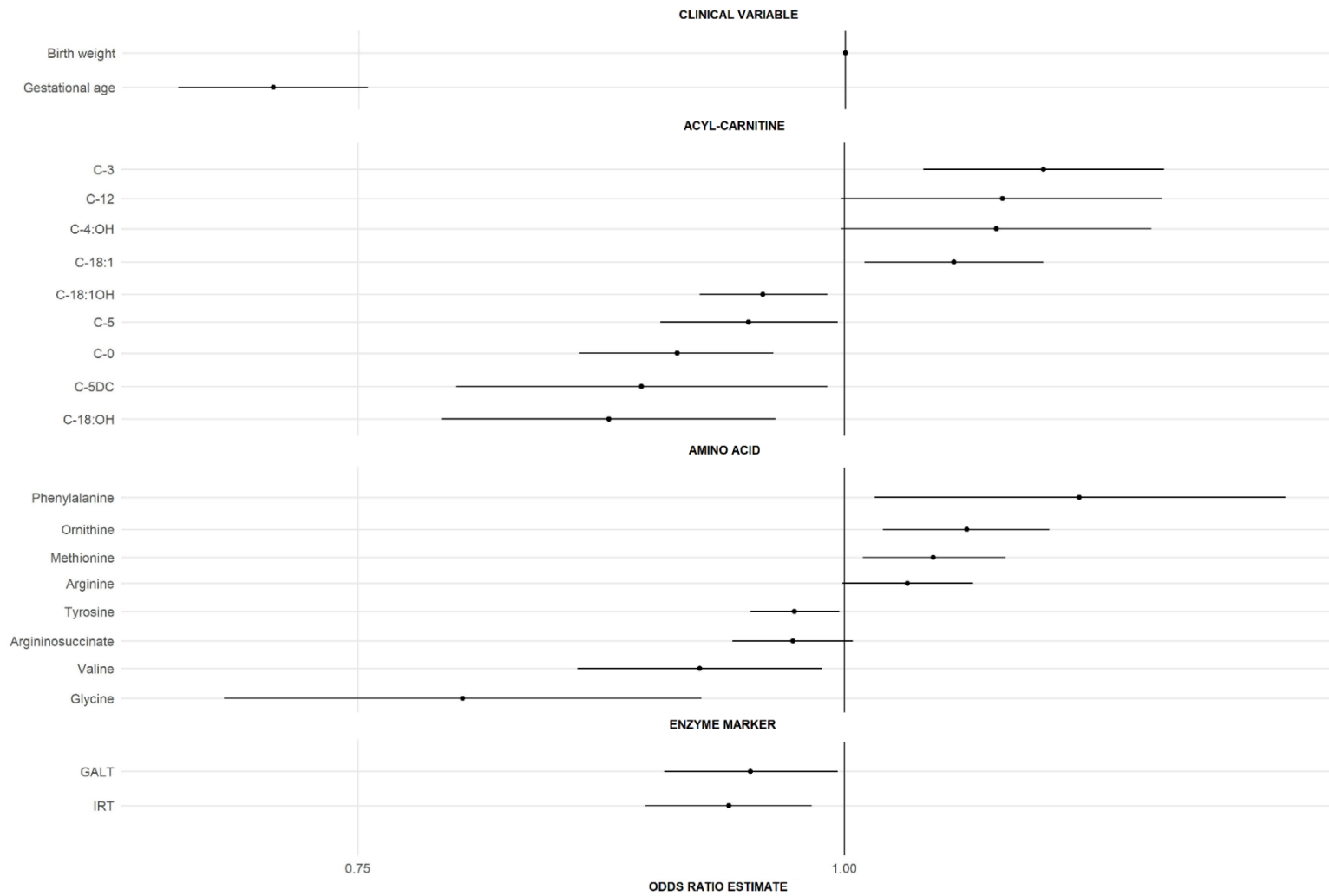
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.2b Analyte ratios in the final model of preterm infants born between 2010 and 2012



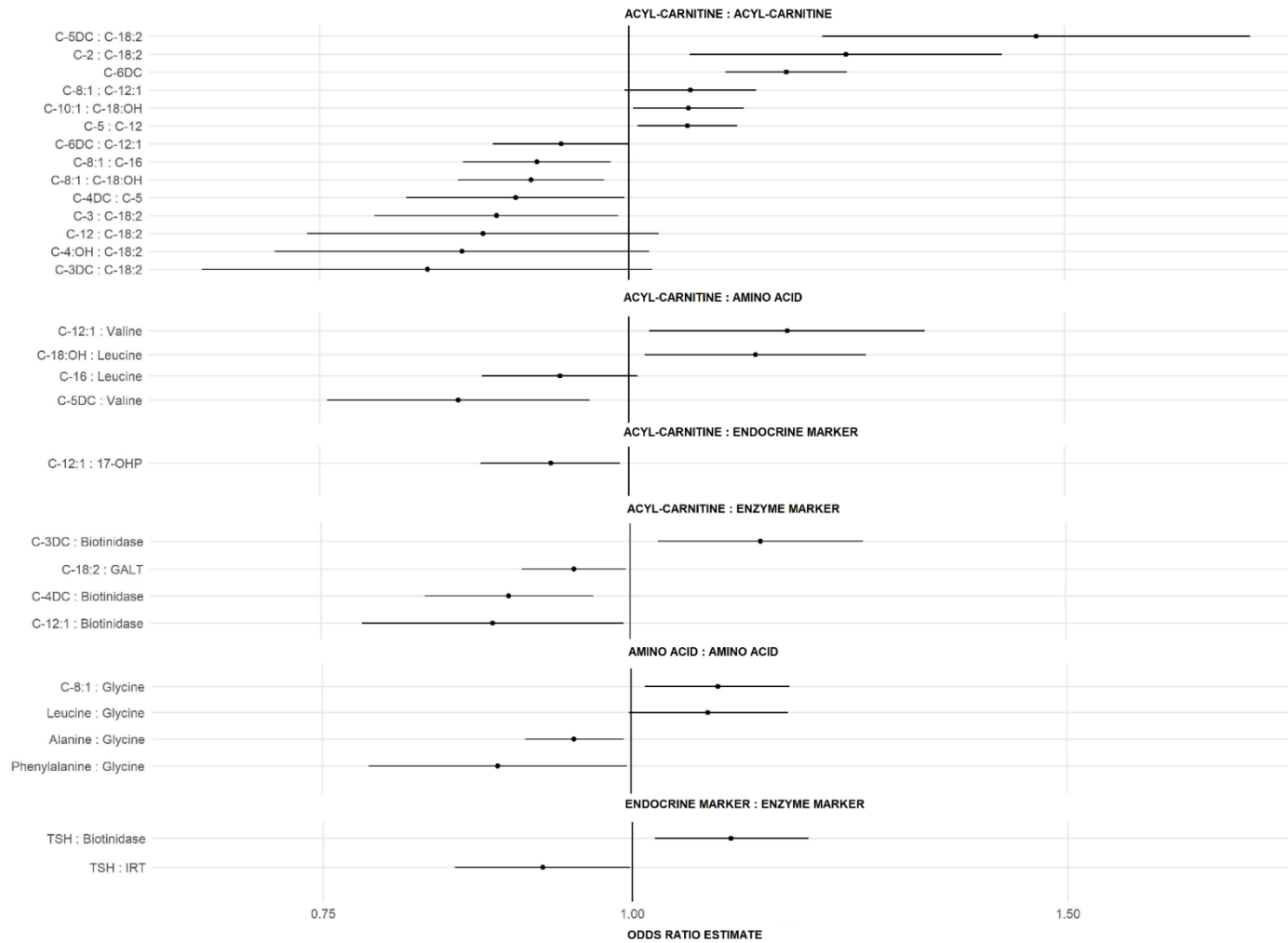
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.3a Individual analytes and clinical variables in the final model of pre-term infants including TPN



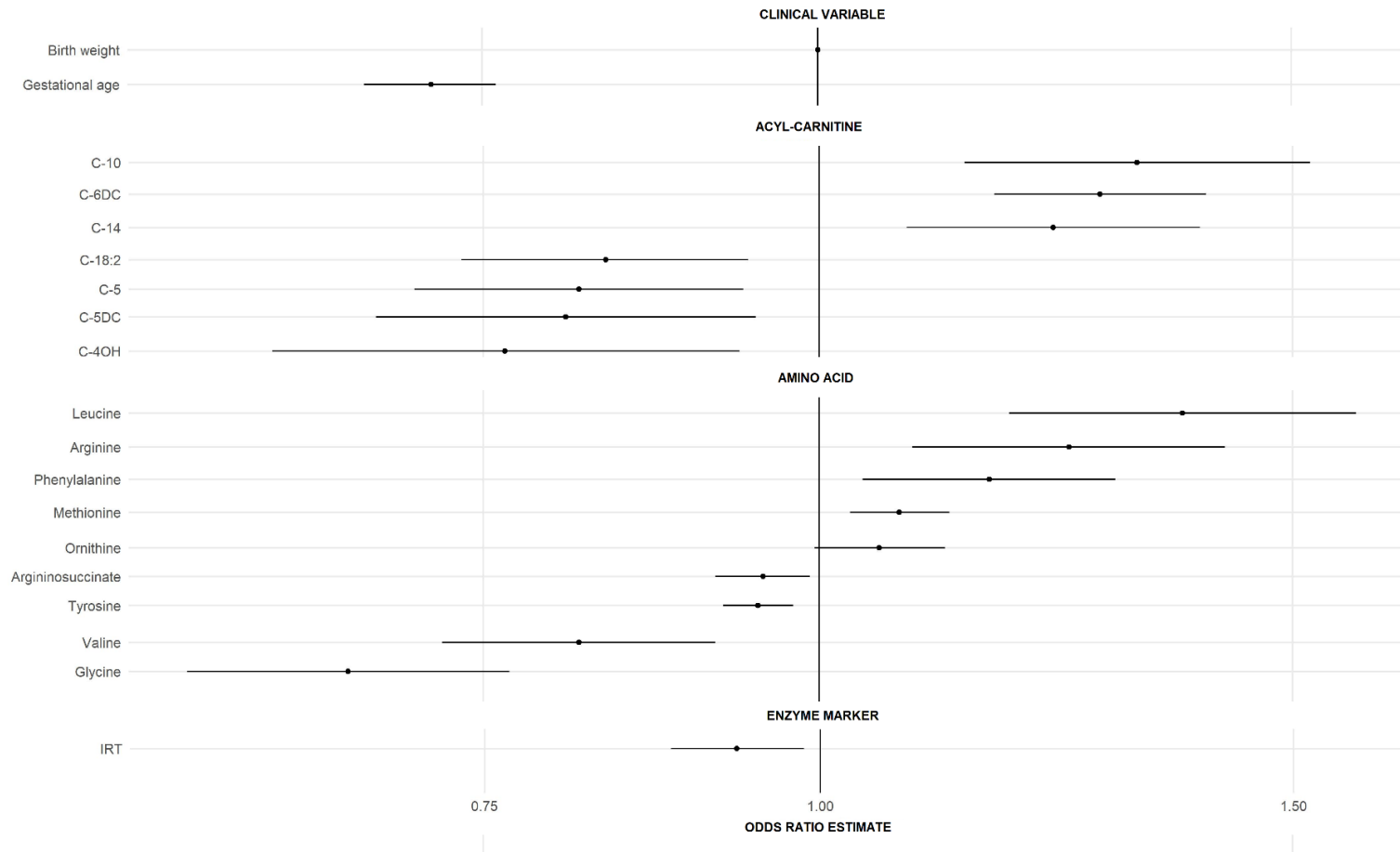
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.3b Analyte ratios in the final model of preterm infants including TPN



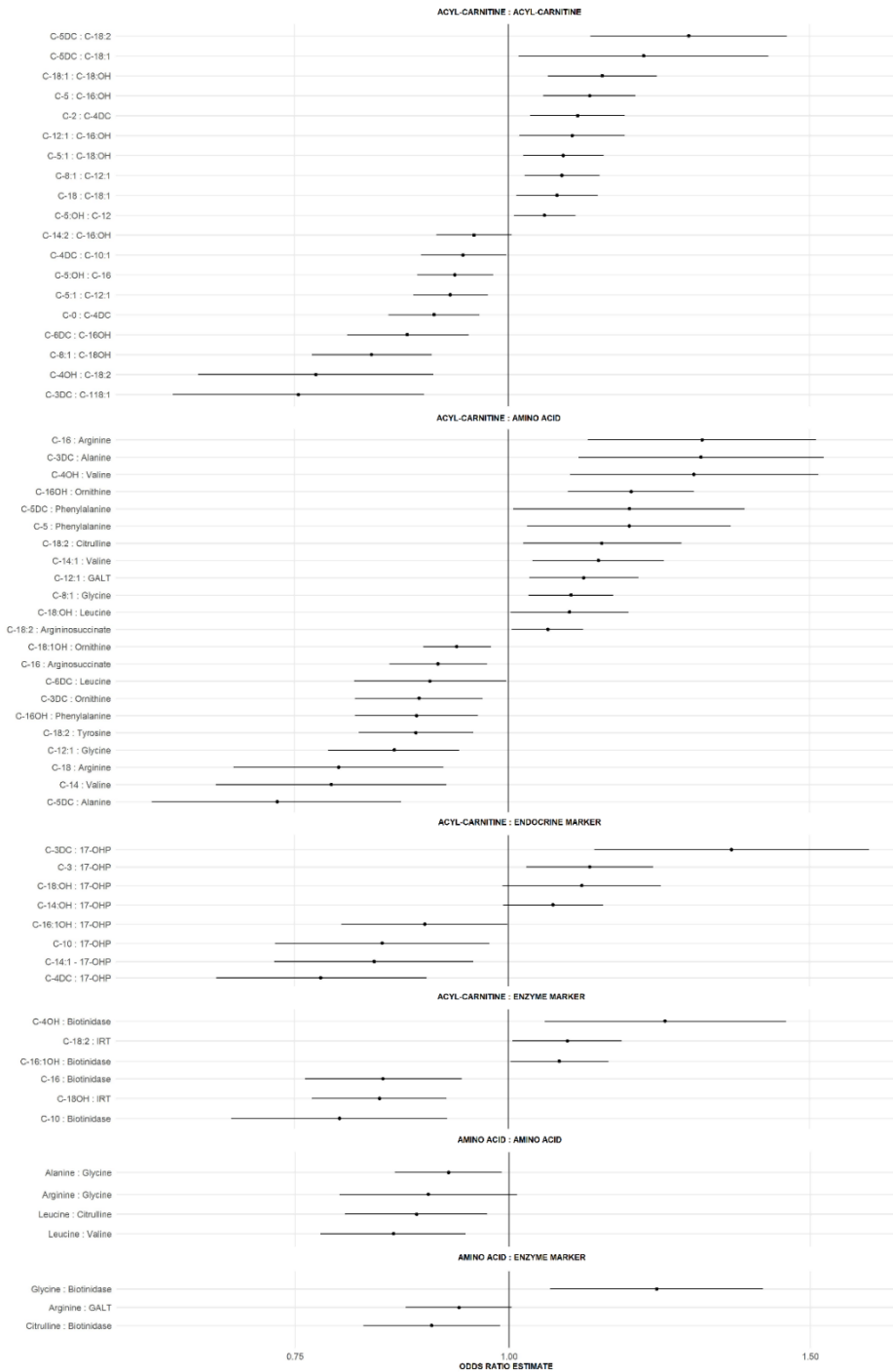
Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.4a Individual analytes and clinical variables in the final model of-term infants with the inclusion of remaining ranked analyte ratios based on Spearman's correlation until the maximum number of parameters was reached



Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.

Supplemental Figure J.4b Analyte ratios in the final model of term infants with the inclusion of remaining ranked analyte ratios based on Spearman's correlation until the maximum number of parameters was reached



Forest plot demonstrating associations of newborn screening analytes with delivery following PPROM, compared with delivery following spontaneous onset of labour. ORs greater than one indicate a positive association with delivery following PPROM, while ORs less than one indicate an inverse association with delivery following PPROM.