Universidade de Aveiro Departamento de Química
Ano de 2019

Mariana Carreira Rodrigues

A Metabolómica na pesquisa de novos biomarcadores para o acompanhamento da saúde dos recém-nascidos prematuros

Metabolomics in search of new biomarkers for health monitoring of preterm newborns

Universidade de Aveiro Departamento de Química
Ano 2019

Mariana Carreira Rodrigues

A Metabolómica na pesquisa de novos biomarcadores para o acompanhamento da saúde dos recém-nascidos prematuros

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, realizada sob a orientação científica da Doutora Ana Maria Pissarra Coelho Gil, Professora Associada com Agregação do Departamento de Química da Universidade de Aveiro.

Esta tese foi realizada no âmbito do projecto CICECO- Instituto de Materias de Aveiro, POCI-01-0145-FEDER-007679 (FCT ref. UID/CTM/50011/2019), financiado por fundos nacionais através de FCT/MCTES. Agradeço à Rede Nacional de RMN (RNRMN), suportada com fundos da FCT.

o júri

presidente

Doutora Maria do Amparo Ferreira Faustino
Professora Auxiliar, Universidade de Aveiro

Doutora Iola Melissa Fernandes Duarte
Equiparada a Investigadora Principal, Universidade de Aveiro

Professora Doutora Ana Pissarra Coelho Gil
Professora Associada C/Agregação, Universidade de Aveiro

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Ana M. Gil for her guidance and all the knowledge she passed on to me. I am thankful to her for giving me the opportunity to pursue this work, for introducing me to the field of metabolomics since that it is new for me and mainly for not letting me give up even when all seemed lost.
I want to acknowledge to Dr. Maria do Céu Almeida, Dr. Fátima Negrão and Nurse Cristina Pita (from Maternidade Bissaya Barreto, Coimbra) for providing the samples and the corresponding clinical information of babies.
I would also like to acknowledge our group for providing some knowledge necessary for successfully conclude this work, particularly to Daniela Duarte for your help with NMR experiments and data analysis.
I am especially grateful to have shared this big journey with Joel for his companionship, support, patience and for giving me the strength to continue.
I am also thankful to my friends for all the support words and patience.
Finally, I would like to thank my parents, for their unconditional love, patience and support.
It was not an easy journey, but in the end it was all worth it.

palavras-chave Parto prematuro, metabolómica, espectroscopia de RMN de protão, urina, biomarcadores

resumo O capítulo 1 divide-se em quatro partes, sendo a primeira parte referente à definição de parto prematuro, identificando os seus principais fatores de riscos e consequentes complicações na saúde a curto e longo prazo. Na parte 2 é apresentada a metodologia aplicada, fazendo-se uma breve introdução aos métodos analíticos e estatísticos usados e sendo ainda referidos os princípios da metodologia usada e estado da arte. A parte 3 descreve a importância dos biomarcadores no diagnóstico, terapia e prognóstico na área da saúde, referindo como são identificados e validados clinicamente. Na parte 4 é descrito o potencial clínico da metabolómica na neonatologia, fazendo referência aos biofluidos usados e aos principais estudos metabolómicos associados ao parto prematuro. Esta parte termina com uma revisão da literatura de estudos realizados em prematuros recém-nascidos.

O capítulo 2 descreve os procedimentos experimentais utilizados para a realização deste trabalho, incluindo apresentação dos grupos, recolha e preparação de amostras, aquisição de amostras e tratamento de dados.

No capítulo 3 é feita a caracterização da composição da urina de recém-nascidos através de espectroscopia de Ressonância Magnética Nuclear (RMN) através da qual foram identificados 49 metabolitos.

O capítulo 4 apresenta o estudo metabolómico de recém-nascidos prematuros a fim de encontrar marcadores relacionados com prematuridade no momento do nascimento. Nesse contexto são apresentadas as variações metabólicas em diferentes estadios de prematuridade (extremo (<28semanas gestacionais), muito prematuro

(28<32 semanas gestacionais) e moderado a tardio (32<37semanas gestacionais)). No capítulo 5 é feito o estudo metabolómico relativamente ao desenvolvimento dos recém-nascidos prematuros durante a sua estadia no hospital até ao momento da alta médica.
Finalmente, no capítulo 6 apresentam-se as principais conclusões deste trabalho, enfatizando o potencial clínico da metabolómica de urina de recém-nascidos prematuros para o acompanhamento destes bebés durante a estadia no hospital, visando o uso de novos biomarcadores indicativos de desvios relativos a complicações de saúde durante esse período.

Keywords Preterm birth, metabolomics, ^{1}H NMR spectroscopy, urine, biomarkers

Abstract

Chapter 1 is divided in 4 parts, comprising a first part referring to the definition of preterm, identifying its main associated risk factors and consequent health complications in short and long terms. In part 2 the applied methodology is presented, giving a brief introduction about the analytical and statistical methods used. The principles of the methodology used and state of the art are also described. Part 3 describes the importance of biomarkers in diagnostics, therapy and prognostics of disease. In part 4 the clinical potential of metabolomics in neonatology is described, referring the biofluids used and the main metabolic studies associated with preterm birth (PTB). This part concludes with a literature review of studies of premature infants and newborns, and premature infants compared to other disorders.

Chapter 2 describes the experimental procedures used to perform this work, including group presentation, sample collection and preparation, sample acquisition and data processing.

Chapter 3 characterizes the composition of newborn's urine using Nuclear Magnetic Resonance (NMR) spectroscopy, which 49 metabolites were identified.

Chapter 4 presents the metabolic study of preterm newborns to find markers related to prematurity at the time of birth. Metabolic variations at different stages of prematurity (extreme preterm (<28gestational weeks), very preterm (28<32 gestational weeks) and moderate to late preterm (32<37gestational weeks)) are presented.

Chapter 5 deals with the metabolomic study of the development of preterm newborns over time, during their stay in hospital.

Finally, the main conclusions of this work are present in Chapter 6, emphasizing the clinical potential of metabolomics in preterm urine for monitoring newborns during their hospital stay, in order to use new biomarkers of deviant behaviours indicative of health complications during that period.

Contents

Abbreviations i
Chapter 1. Introduction 1
1.1. Premature birth (PTB) and effects on infant health and development 1
1.1.1. Development of the newborn and risk factors of preterm birth 1
1.1.2. Frequent health complications of PTB 3
1.2. Metabolomics: concept, strategies and application in disease research 7
1.2.1. Concepts and definitions 7
1.2.2. Analytical strategies in metabolomics 9
1.2.2.1. Nuclear Magnetic Resonance (NMR) spectroscopy 9
1.2.2.2. Mass spectrometry (MS) 11
1.2.2.3. Pre-processing and statistical analysis of data 11
Principal component analysis (PCA) 12
Partial least squares discriminant analysis (PLS-DA) 13
1.3. Metabolomics in search of disease biomarkers: clinical importance, identification and interpretation 13
1.4. Metabolomics to monitor the health of premature newborns 15
1.4.1. The clinical potential of metabolomics in neonatology 15
1.4.2. Application of biofluid metabolomics to premature newborns 17
1.4.2.1. Premature compared to healthy newborns 18
1.4.2.2. Premature compared to newborns affected by other disorders 20
1.5. Aims of thesis 22
Chapter 2. Experimental Section 23
2.1. Sampling 23
2.1.1. Sample and metadata collection 23
2.1.2. Definition of sample groups 23
2.2. NMR spectroscopy 25
2.2.1. Sample preparation for NMR analysis 25
2.2.2. Acquisition and processing of NMR data 26
2.2.3. Pre-processing and multivariate analysis of NMR data 26
Chapter 3. Typical NMR spectra of newborn urine and assignments 27
Chapter 4. Cross sectional study of newborns with different development degrees at birth – Study 1 33
4.1. Visual comparison of spectra per class 33
4.2. PCA and PLS-DA analysis of prematurity compare to controls 35
4.2.1. Proposed metabolic interpretation of changes related with prematurity at birth time 39
4.3. Statistical analysis of newborn urine spectra at birth (without outliers) 44
4.3.1. Extremely Preterm + Very Preterm vs Controls 44
4.3.2. Extremely Preterm/Very Preterm vs Moderate to Late Preterm 48
4.3.3. Moderate to late Preterm vs Controls 51
4.3.4. Effects of different stages of prematurity on urine composition: proposed metabolic interpretation 54
Chapter 5. Longitudinal study of individual newborns with different prematurity degrees – Study 2 57
5.1. Statistical analysis of newborns urine spectra and metabolites trajectories 57

5.2. Proposed metabolic interpretation of changes during premature newborns development until term time ... 66
Chapter 6. Conclusion and future perspectives ... 67
Bibliography ... 69

Abbreviations

2D	Two dimensional
EPT	Extremely preterm
g.a.	Gestational age
g.w.	Gestational week
PTB	Preterm birth
PCA	Principal Component analysis
PLS-DA	Partial least squares regression discriminant analysis
PQN	Probabilistic quotient normalization
MLPT	Moderate to late preterm
MVA	Multivariate analysis
MB	Maternal blood
MU	Maternal urine
NMR	Nuclear magnetic resonance
NB	Newborn
NU	Newborn urine
UCB	Umbilical cord blood
VPT	Very preterm

Chapter 1. Introduction

1.1. Premature birth (PTB) and effects on infant health and development

1.1.1. Development of the newborn and risk factors of preterm birth

A human full-pregnancy has the duration of 37 to 42 gestational weeks (g.w.) since the first day of the woman's last menstrual period and comprising three trimesters. These stages are responsible for fetal development because each one is characterized by specific hallmarks. Preterm birth (PTB) is defined by the World Health Organization (WHO) as any birth occurring before 37 completed g.w. or 259 days of gestation and is considered the leading cause of neonatal death in children under 5 years of age worldwide (1,2). Based on gestational age at birth, PTB is subdivided into extremely preterm (<28 g.w.), very preterm (28<32 g.w.) and moderate to late preterm (32<37 g.w.) (3).

Each year 1 in 10 babies are born prematurely, which corresponds to an estimated 15 million that are born too early, representing a preterm birth rate of 11.1% (3). The rate of premature births has increased in many countries over the years, especially those in development, where the rates are highest (11.8%) compared to lower middle-income countries (11.3%) and high-income countries (9.4%) (3). In *Figure 1 a)* a distribution of the number of preterm births by gestational age and region is represented. In relation to gestational age it is noted that a high rate of PTBs occurs between 32 and 37 g.w. (moderate to late preterm) and much lower rates characterize births occurring before 28 g.w. (extremely preterm birth) and between 28 and 32 g.w. (very preterm birth). The continents with the highest absolute number of PTBs are Africa and Asia, with an average of 12.8% per year. This can be explained not only by the high number of births, but also by the lack of feasible and cost-effective care necessary to monitor pregnancies and prevent these births (such as ultrasound equipment), but also to treat babies that need special care to survive, especially basic care for infections and breathing difficulties. Although preterm rates are lower in higher-income countries, in the United States almost 1 in 8 babies is born prematurely (12% of total births) (3). This high rate shows that even in a developed country, there are possibilities of multiple pregnancies, increasing the pregnancies among women at higher ages, greater manipulation of assisted techniques and increase use of practices such as induction of labor and cesarean delivery (4). Globally, mortality increases with decreasing of gestational age (3).

In Portugal, according to the *Instituto Nacional de Estatística (INE)*, the rate of PTB increased slightly from 7,4% to 7,8% since 2011 until 2016 (*Figure 1 b)*) (5).

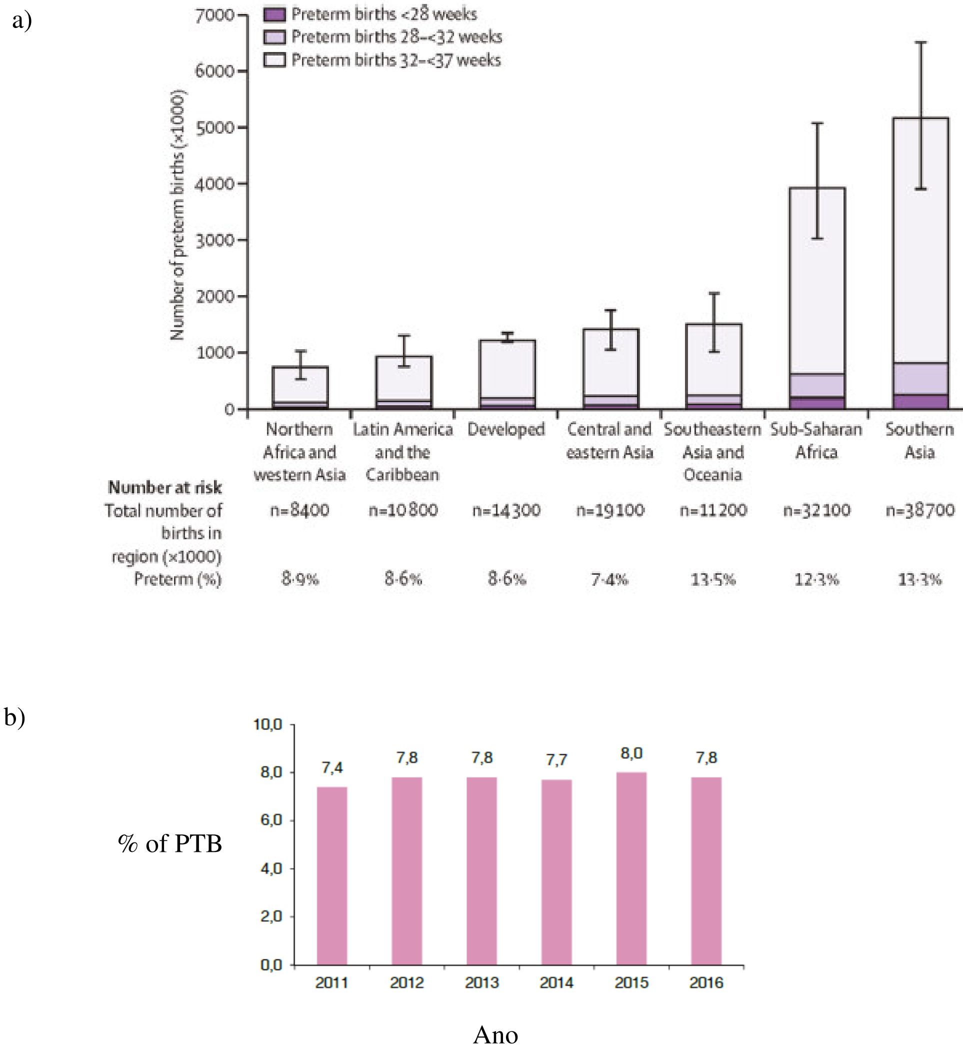


Figure 1 – a) - Distribution of the number of PTB by gestational age and region [Adapted from (3)]; b) rate of PTB in Portugal since 2011 to 2016 [Adapted from Dados Demográficos 2016 do Instituto Nacional de Estatistica (INE) (5)].

The obstetric origins leading to PTB can be divided in three subgroups: spontaneous preterm birth with intact membranes (45% of cases); preterm rupture of membranes (PROM) (25% of cases) and indicated delivery of a premature infant for both benefits of the mother and the infant in which labor is either induced or the infant is delivered by caesarean (30%

of cases) (6). These clinical subgroups can explain most of preterm births (1), however, no specific cause for preterm birth may be identified because of its multifactorial nature, which means there are many risk factors that may lead to the condition (7). Some of these risk factors have been associated with the maternal medical problems such as pre-eclampsia, hypertension, gestational diabetes mellitus, diabetes mellitus, infections, placental or uterine abnormalities, gestational haemorrhages; others are associated with maternal pregnancy history such as multiple gestations, history of abortion, multiple 2nd trimester spontaneous abortion, history of infertility; or with maternal age, socioeconomic status and lifestyle factors such as smoking, consumption of alcohol or drugs consumption (8,9).

1.1.2. Frequent health complications of PTB

The significance of PTB lies in the complications caused by prematurity and their impact on the infant's survival and development (9). The complications of PTB result from the immaturity of organs since premature newborns have absence of the complete formation of certain organ systems. Naturally, the lower the gestational age, the greater the immaturity of the organs and, consequently, the greater the risk of acute of neonatal illnesses. Concomitantly, this newborn is unprepared for extra-uterine environment, which leads to various complications soon after birth (9) and, consequently, more necessity of life support. *Figure 2* shows a representation of the most critical periods of a normal fetal development, that occurs between the 3rd and 16th g.w. and is possible to observe that the high rate of anomalies occurs in central nervous system (CNS), heart, limbs, ears, eyes, teeth, palate and external genitalia. A wide range of organ systems is affected (*Table 1*) by complications associated to PTB (9)). Noteworthy, lungs are the last organs to develop, so respiratory problems are the most frequent complications in premature newborns. Nevertheless, gastrointestinal, immune and cardiovascular systems are affected as well (9).

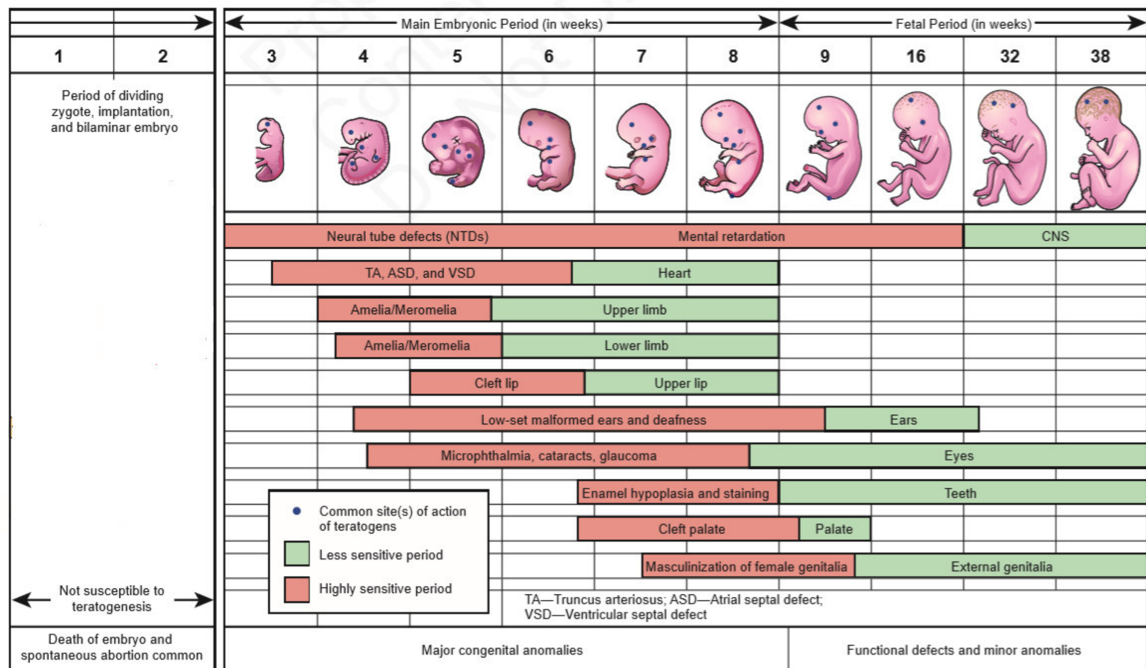


Figure 2 **-** Representation of the critical periods in prenatal fetal development [Adapted from (10)].

Table 1 **-** Health complications of PTB in terms of fetal development and during the perinatal and neonatal periods [Adapted from (9)]. Abbreviations: L. and R.S.- Lung and Respiratory System; C.N.S.- Central Nervous System; G.I.S- Gastrointestinal System; I.S.- Immune System; H.S.- Hematologic System; C.S.- Cardiovascular System.

Health Complication	**System**
Respiratory Distress Syndrome (RDS)	L. and R.S.
Bronchopulmonary Dysplasia (BDP) / Chronic Lung Disease (CLD)	L. and R.S.
Apnea	L. and R.S.
Intraventricular haemorrhage (IVH)	C.N.S
Cerebral Palsy	C.N.S
Necrotizing enterocolitis (NEC)	G.I. S.
Sepsis	G.I. S.
Infection	Skin and I. S.
Patent ductus arteriosus	C.S.

Lung and Respiratory System complications

Gas exchange is the primary function of the lung and the breathing of amniotic fluid in and out is essential for initial lung development. Alterations in breathing or lack of amniotic fluid can lead to pulmonary hypoplasia (undeveloped lungs). In the normal development of the lungs an important complex mixture of lipids and proteins (surfactant) is produced at approximately 30 to 32 g.w.. Infants that are born before 30 g.w. can have deficits in surfactants and consequently, disruptions in normal lung maturation, leading to short and long-term lung and respiratory defects (9,11). However, certain substances such as corticosteroids can accelerate the formation of surfactants and in case of situations such as prolonged membrane rupture and intrauterine growth retardation it is possible to accelerate surfactants formation; contrary, in case of situations such as Diabetes mellitus and hyperglycemia from the mother retard surfactant formation (11).

Premature infants can acquire two principal injuries related to lung: Respiratory Distress Syndrome (RDS) and Bronchopulmonary Dysplasia (BPD) (also called, Chronic Lung Disease (CLD)) (9). Respiratory Distress Syndrome (RDS) is an acute lung injury caused by deficits in surfactants and affects mostly the extremely PTB (9). RDS complications include pneumothorax, intraventricular hemorrhage, BDP or even death (9). Bronchopulmonary Dysplasia (BPD) is the most common long-term complication of prematurity (12) and consists of oxygen dependence at 36 weeks since the last menstrual period of mother and for at least 28 days from birth (9,12). BDP is characterized by injuries on alveolar and vascular development (12). Infants with BPD may be frequently diagnosed with asthma and suffer constant wheeze (12) and develop respiratory infections with possibility of rehospitalizations, exogenous surfactant or a ventilator (9).

Central Nervous System complications

The final process of CNS development begins at 6 months of gestation and continues during the childhood (9,13). CNS organization is crucial for the development of several areas of the brain (13). CNS can be damaged if not fully developed and, in that case, preterm newborns are more vulnerable to CNS injuries. Thus, between 23 and 26 g.w. there is a risk of intraventricular hemorrhage (IVH), cerebral hemorrhage (CH) and periventricular leukomalacia (PVL). Many factors contribute to brain injury, such as high blood obstruction, hypotension, poor autoregulation of cerebral blood flow, hypoxia and harmful inflammatory

substances (9). Additionally, infants that are born premature and suffer from IVH, CH or PVL are at an increased risk of neurodevelopmental disabilities at long-term including, cognitive/behavioral deficits, visual, perceptual problems and high risk to develop cerebral palsy (1,9,13).

Gastrointestinal System complications

The gastrointestinal tract digests and absorbs food and nutrients and has immune and endocrine functions (9). The mouth, esophagus, stomach and intestines are part of the gastrointestinal tract and this system starts to form in the fourth g.w. (9). At 20 g.w. the stomach and the intestines should be completely formed (9).

During and after birth the newborn's gastrointestinal tract is colonized by microbes from the mother and the environment.

As a consequence of premature birth, several complications may occur in the gastrointestinal system, compromising its normal functioning. The most common PTB complications are feeding intolerance and Necrotizing enterocolitis (NEC) (9,14). Feeding intolerance, is the incapacity of the gastroinstestinal (GI) tract to digest food due to its immaturity (9). Furthermore, in case of very immature and sick infants, they must receive parenteral (intravenous) nutrition in order to acquire amino acids, glucose, electrolytes and lipids necessary to normal growth and development (9). NEC is another PTB complication, which consists of acute injury of the intestines which cause inflammation (9). NEC occurs frequently within 2 weeks after birth and is characterized by feeding difficulties, abdominal swelling, hypotension, and signs of sepsis (9). Moreover, preterm infants that go through GI complications can have short-term and long-term morbidities which include malabsorption (cause by removal of large portions of the intestine), repeated surgical procedures, liver failure, prolonged parenteral nutrition, poor nutrition and rehospitalization (9).

Infections and Immune System complications

The immune system of preterm infants is extremely vulnerable to infectious agents (such as bacteria, viruses and other risk organisms) since it is not prepared for extrauterine life (9,15). These infections can either occur during pregnancy (through abnormalities occurring between placenta and fetus), delivery (through the mother's bacteria) or due to the extrauterine environment (through undeveloped skin, lung or GI tract) (9). Besides that, an elevated rate of infections may also occur in hospital. Pneumonia, sepsis, meningitis and

urinary tract infections are the most common consequences of infections with these agents on preterm infants (9). The immune system maturation occurs in the postnatal period and may vary with the infant's diet (e.g. breast milk) and environment (e.g. pathogens and allergens) (15).

Cardiovascular System complications

The fetal heart is completely formed at the end of the 6th gestational week. After birth ductus arterious may close, but in preterm newborn that may not occur, resulting in a patent (open) ductus arteriosus. Consequently, heart failure and reduced blood flow may occur, bringing complications, such as IVH (intraventricular hemorrhage), NEC (Necrotizing enterocolitis), BDP/CLD (Bronchopulmonary Dysplasia/ Chronic Lung Disease) and death if not treated (9).

Importantly, prematurity and its associated complications cause psychological and emotional demands to parents and family. Additionally, premature newborns can spend long periods in the hospital receiving specially care leading to excessive economic expenses. So, it is important to carry out more research to find different strategies to alleviate these problems.

1.2. Metabolomics: concept, strategies and application in disease research

1.2.1. Concepts and definitions

"Omic" sciences seek to understand organisms' cellular function in living systems and their biological alterations upon different stimuli (16). "Omic" sciences comprises genomics, transcriptomics, proteomics and metabolomics which study genes, RNA, proteins and metabolites, respectively (16) (*Figure 3*). Metabolomics is the more recent "omic" and it is considered by many authors as the most promising. It is based on the study of all metabolites (called metabolome) present in a biological system (e.g. biofluids, cells, organisms) (17,18). Metabolites are molecules with less than 1kDa and because they represent the downstream expression of genome, transcriptome and proteome, they can closely reflect the phenotype of an organism at a specific time. Also, they can be endogenous if they are intermediaries or end products of metabolism or exogenous when a cell or organism ingest or synthesize some substances arising from foods or drugs. Metabolites includes amino acids, organic

acids, fatty acids, nucleic acids, lipids, hormones, drugs, peptides, vitamins, food additives, among others) (19).

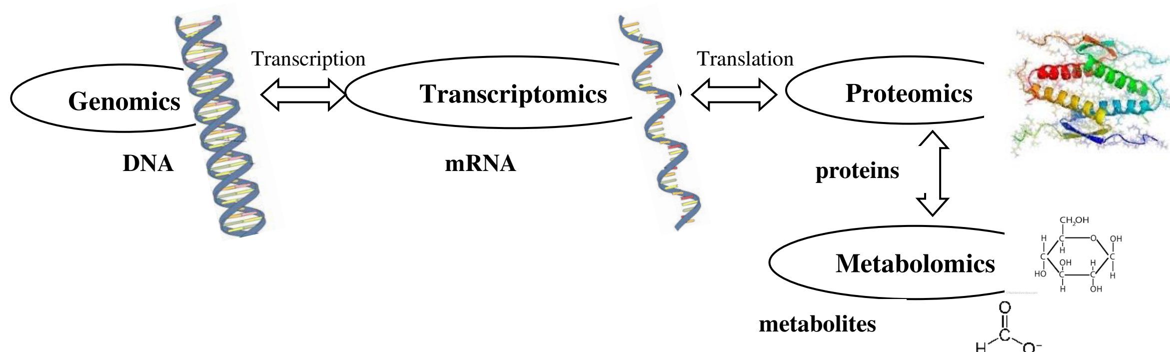


Figure 3 – Schematic representation of the "omics" technologies.

These metabolites contain most of the information necessary to provide the metabolic state of cell, tissue or organism when exposed to genetic variations, drug treatments or pathophysiological variables (17,20).

Metabolomics has grown a lot due to its wide range of applications, namely in nutrition, biology, drug discovery and human disease (e.g. cancer, cardiovascular diseases, neurological diseases, diabetes, respiratory diseases, gastrointestinal diseases, among others) (17). The biochemical field, including the search for diagnostic and prognostic biomarkers, is one of the most developed areas. In addition to biomarkers, metabolic predictors for treatment response are also investigated (17).

The general steps performed in metabolomic studies are illustrated in *Figure 4* and are discussed in the next sections. Metabolomic studies comprise the analysis of biological fluids (this topic is detailed for PTB in Section 1.4) (e.g. serum, plasma, urine, saliva, blood, umbilical cord blood, cerebrospinal fluid, amniotic fluid), tissues and cell extracts, the profile or metabolic alterations being studied over time, when exposed to various changes (e.g., extrauterine environment). Metabolomic analysis is generally built as case-control studies, comparing groups of samples coming from individuals with different pathophysiological states.

The most used analytical methods in metabolomics are Nuclear Magnetic Resonance (NMR) spectroscopy and Mass spectrometry (MS) (21,22). Fourier-transform infrared spectroscopy (FTIR) is less frequently used for these studies, however it has a lower cost

(16). Both NMR and MS methods have high resolution to detect molecular species separately (23), thus having the ability to generate complex datasets. The results obtained with these analytical strategies are commonly vast and require pre-processing and data mining methodologies based on multivariate analysis (MVA, described in the section 2.2.3) for their handling and interpretation in terms of potential metabolite biomarkers.

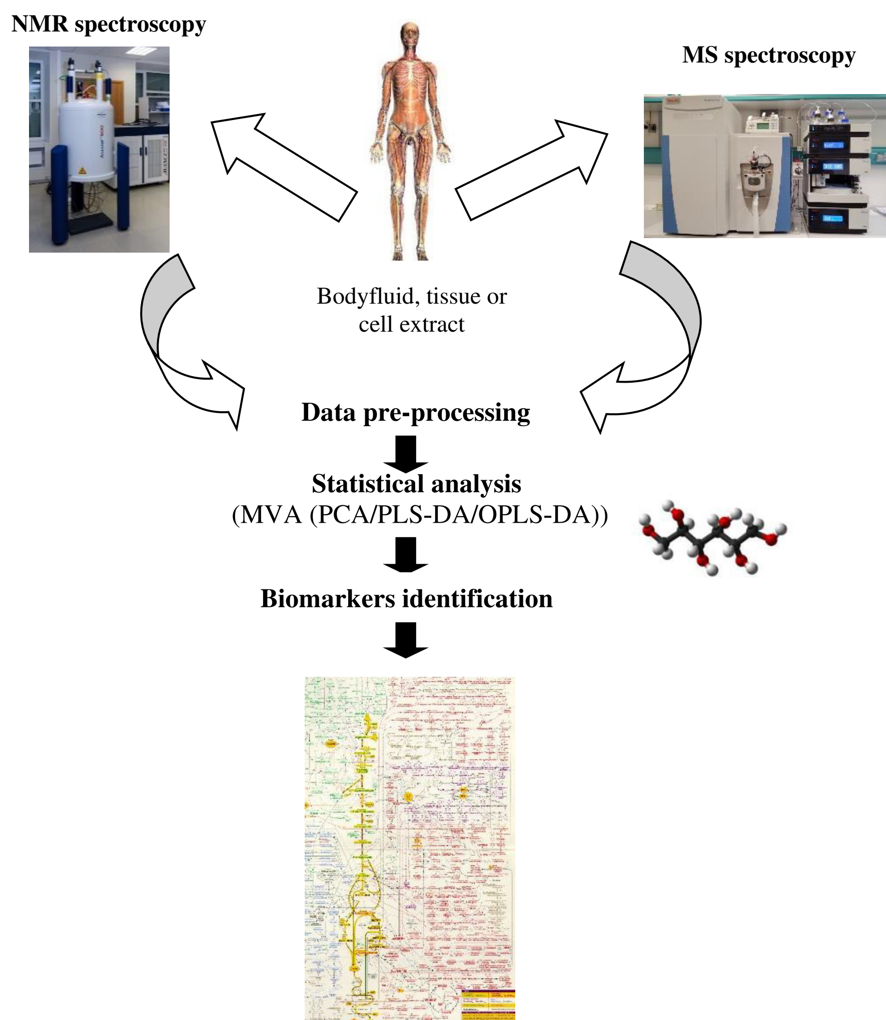


Figure 4 – Illustrative scheme of general steps involved in metabolomic studies.

1.2.2. Analytical strategies in metabolomics

1.2.2.1. Nuclear Magnetic Resonance (NMR) spectroscopy

NMR is a technique extremely reproducible, robust, quantitative and requiring easy manipulation of sample (16,19,24,25). NMR can be applied to either liquid and solid samples

(16). It is a rapid technique, available to study several nuclei (e.g.^{1}H, ^{13}C) (19) and allows the quantification of several metabolites at the same time (24). However, it has low sensitivity. It requires skilled technicians and is relatively expensive (16,19).

For clinical studies, ^{1}H NMR is widely exploited because the majority of known metabolites in biological samples contain hydrogen. NMR spectroscopy is based on the magnetic properties of some atomic nuclei (20) where its energy absorption and release occur (17). In fact, protons have a rotating movement, called nuclear spin, associated to an angular momentum P (26), leading atoms to behave like magnets (27). Atoms are surrounded by electrons that modify the magnetic field because they may be positioned at different sites on the molecule. The modified magnetic field causes a shielding effect in the nucleus (28). Thus, according to their chemical environment, each nucleus will resonate at different frequencies and consequently, the positions in the NMR spectrum will be different (27,28). This position is designated as chemical shift (δ), expressed in ppm (parts per million) and mathematically is given by *Equation 1*, where ν is the NMR frequency machine and νref is the frequency of the peak of the reference compound (used for calibration of NMR spectra) (δ=0) (27,29):

$$\delta = 10^{6} \times \left(\frac{\nu - \nu\mathrm{ref}}{\nu\mathrm{ref}}\right)$$

Equation. 1

The most common compounds used as chemical shift references in both ^{1}H and ^{13}C spectra are trimethylsilylpropanoic acid (TSP) and tetramethylsilane (TMS). The scale of chemical shifts in ^{1}H NMR spectra varies from 0 to 14 ppm and the peak appears in different positions of this scale consoant the functional group within a molecule. The peak area allows metabolites quantification since it is proportional to the number of protons and therefore proportional to the concentration of the referred molecule (30).

For consistent results several factors should be considered such us identical NMR tubes, same instrumental parameters (temperature, shimming, pulse sequences, acquisition times and data-processing parameters) (24). Many signals can be directly annotated based on their chemical shifts, J coupling constant and the multiplicity of the resonance peaks using knowledge from international databases (e.g. Human Metabolome Databese (HMDB), ChenomX, Biorefcode AMIX, Bruker). However, the assignment of NMR spectra of a mixture can be a complex procedure and it may not be possible to identify some of the peaks. For this purpose, two-dimensional (2D) NMR spectroscopy increases signal dispersion and

elucidates the connectivity between nuclei (17), providing additionally information and subsequently helping to identify further signals (23). The 2D methods mostly used for assignment are Correlation Spectroscopy (COSY), Nuclear Overhauser Effect Spectroscopy (NOESY), Heteronuclear Single Quantum Correlation (HSQC), Heteronuclear Multiple Bond Correlation (HMBC) and (Total Correlation Spectroscopy (TOCSY) (31).

1.2.2.2. Mass spectrometry (MS)

MS is the technique most used in metabolomics, featuring high sensitivity, selective and providing rapid analysis (16). Mass spectroscopy is a technique based on the ionization of a molecule where all the ions are separated according to their mass to charge ratio (m/z) and consequently your intensity is proportional to their relative abundance (16). Ions that are observed in the mass spectrum can elucidate the nature and chemical structures of their precursor molecules. Under ionization, there is a loss of an electron of the molecule and this cationic radical formed can be fragmented generating new ions, called product ions, and therefore provide the specific structural information about the molecule (32). Is important that spectrometers function under vacuum to prevent that the ions collide with other gaseous molecules, that if they existed, they would cause unwanted reactions and the complexity of the spectrum will increase.

MS may need to be coupled with analytical separation techniques, usually liquid chromatography (LC) and gas chromatography (GC) (16). (Liquid chromatography coupled with MS (LC-MS) consists in combination of HPLC and MS, making LC-MS an embracing technique due to high variety of stationary phases and different separation methods (16), however require long analytical time per sample. On the other hand, Gas chromatography coupled with MS (GC-MS) has high accuracy and repeatability of results, require small samples amounts and has high discrimination of molecules with a very similar structure (19). Both LC-MS and GC-MS require long sample preparation, are destructive, and the equipment used is moderately expensive (19), nonetheless the higher sensitivity of MS makes it more advantageous for several studies.

1.2.2.3. Pre-processing and statistical analysis of data

Both NMR and MS-based techniques generate large and complex dataset and therefore, must be first pre-processed for multivariate modeling (33). Data shows variations in the alignment of the peaks and intensities which can be caused by the differences between

samples (i.e., pH, temperature, shim or spectral calibration) (34). Thus, these variations will be corrected by pre-processing steps such as alignment, normalization and scaling. In NMR spectra variations in chemical shifts which will cause variations in positions of peaks with same metabolic features. So, to correct that shift, alignment will be applied. Since that biological fluids are applied in that analysis, it is important to remove certain tendencies to obtain only biological differences in data. Normalization is applied to remove these biases to be possible an accurate quantification of the characteristics in a metabolomics analysis. A variance in the average metabolite's abundance found in biological samples can cause an obstruction of the important biomarkers. Thus, scaling is applied, dividing each variable by a scaling factor (35).

In metabolomics the most typical statistical method used for deal with this high data information is multivariate statistical analysis (MVA) (36). The most common MVA methods in metabolomics are principal component analysis (PCA) (unsupervised), partial least squares discriminant analysis (PLS-DA) (supervised) and orthogonal partial least squares discriminant analysis (OPLS-DA) (supervised) (36,37).

Principal component analysis (PCA)

The aim of unsupervised methods such as PCA is to reduce the excess of information (decreasing complexity) and produce possible grouping of samples (33). Thus, PCA is a technique that uses orthogonal transformation to convert the information in the original set of correlated variables into a set of linearly uncorrelated variables, called principal components (PC) (17,38). PCA constructs a two-dimensional data table, known as matrix X, where variation is reduced (39). The plot formed in this model is made by the relationship between observations or samples in the plane, represented as points. The scores plot represents the observations or samples and the position of each object can be detected and related to each other. Thus, if points are positioned closer to each other have similar profiles, while if they are distanced from each other they are more distinct (39,40). The loadings plots represent the relationship between the measured variables and the contribution of each variable of the sample (39,40). A central characteristic is that directions in the scores plot correspond to directions in the loadings plot (39). Concluding, it is important to refer that the scores plot shows the observations that are similar/dissimilar, and the loadings plot show which are the variables that contribute more for this similarity/dissimilarity between observations (39).

Partial least squares discriminant analysis (PLS-DA)

PLS is a technique that measures the quantitative relationship between two matrices X and Y which contains analytical data and quantitative values (e.g. concentration of an endogenous metabolite) respectively (39). This technique can be used in discriminant analysis, known as PLS-DA. PLS-DA is a technique that projects the data into new space (low-dimensional space), showing the association between samples and variables. Additionally, this is a latent variable method where the projection refers anteriorly by decomposed variation into latent variables (LV), which are linear combinations of the original variables (40). The first LV in a PLS-DA model contains the highest variance between classes (40).

The results obtained from supervised multivariate analysis are analyzed through the significance level of metabolites that contribute physiological or pathological state under study. Then these metabolite peaks are integrated and various univariate tests (student t-test) and Python program are applied for confirm the statically significance of individual biomarkers and tests correction to compensate the false discovery rate (such as Bonferroni method and Monte Carlo test).

1.3. Metabolomics in search of disease biomarkers: clinical importance, identification and interpretation

Biomarkers are defined by the National Institute of Health (NIH) as: "*Biological characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathological processes or pharmacological responses to a therapeutic intervention.*" cited by Biomarkers Definition Working Group (41). Therefore, biomarkers correspond to a specific metabolic deviation corresponding to a biologic perturbation caused by a disease, a change in biological state, or an environment exposure (42,43). Biomarkers are extremely used in a high range of diseases, such as in diagnostic of markers for diabetes, heart failure and prostate cancer; markers of prognosis of heart failure, ovarian cancer and colorectal cancer (44). Thus, in the clinic biomarkers are extremely important for diagnosis, prognosis, prediction, monitoring and measuring biological responses of treatments or other external stimuli (24).

PTB born children carry much greater risk to develop some diseases later in life. So, for reduce this high rate of mortality and morbidity, the search for new predictors as biomarkers,

have increased (45). Because of that the identification of biomarkers is considered one of the most promising applications of metabolomics sciences (17).

In neonatal phase, NB has rapid alterations, causing a metabolic change and because of that some doubts have appear in relation to biomarkers for neonatal diseases (45). Thus, when doing a study for a discovery of a new biomarker it is important to test the utility of the used models in real clinical settings. Performance assessment (such as sensitivity and specificity) and model validation are used for this evaluation. In this performance assessment the proprieties can be measured by receiver operating characteristic (ROC) curves, which consist in a comparison of specificity and sensitivity according to a specific decision boundary. They are frequently summarized by the area under the ROC curve (AUC metric), that classifies a biomarker as useful (AUC>0,7) or useless (AUC<0,5) for clinical (17). So, it is important to find a metabolic biomarker to classify diseases with as high specificity and sensitivity as possible (46). In relation to validation methods two of the methods used are: permutation testing and cross-validation. The permutation testing consists in random permutation of classes of sample groups a coupled time, calculating the statistic in each dataset. The obtained results must be outside of the confidence intervals (95 or 99%) of estimated null distributions (H0) to be considered significant (17). On the other hand, cross-validation consists in subdividing the data into a training set and a test set in different segments (47). Other validation method which is frequently referred in PLS-DA plots is Q^2 value and indicates the validity of discrimination (48). Q^2 is the prediction power and its ideal value is 1.

In ^{1}H-NMR based studies the identification is made overlapping the measured NMR peaks with a reference metabolite, that can be found in metabolite spectral databases such as HMDB (Human Metabolome Database) and 2D already mentioned. Through the analysis of metabolites existing in the biological sample analysed and through the biological meaning, it is possible to regroup these metabolites according to the same biological process and to the direct or indirect connectivity by one multiple enzymatic reaction. Databases such as KEGG (Kyoto Encyclopaedia of Genes and Genomes) can help to the researchers with vast information of metabolic pathways (17).

Advances in the search of metabolic biomarkers have increased over the last years, which together with the development in the analytical techniques of NMR and MS (37), have

contributed to the growth of metabolomic applications in health research, including children health (42).

1.4. Metabolomics to monitor the health of premature newborns

1.4.1. The clinical potential of metabolomics in neonatology

The neonatal phase corresponds to the first 28 days of life (37). The first days of life of a newborn are precious not only because the lower the gestational age, the greater the necessary care, but also for the detection of factors that can influence preterm birth or possible related morbidities (49). A way of reducing the neonatal mortality rate is focused on search of early biomarkers capable of predicting the risk of developing diseases and monitoring the route of disease in cases when diseases that already exist or simply for monitoring when is detected early (45). Thus, in this area it is important to adopt “omics” strategies, specially the metabolomics because the metabolites are the trustiest indicators of phenotypes. So, changes occurring in metabolomics are considered more significant when compared with other omics (33,37,46).

The clinical potential of metabolomics in neonatology has increased as evidenced by a growth in publication numbers (30,42,50), demonstrating how physiological variables or pathological conditions may change fetal development or/and their adaptation to extrauterine environment, principally in neonatal phase (50,51). These changes are studied by the alteration on metabolites which may allow improved diagnosis, therapies and prognosis between early life and later disease development (37). For example, the development or appearance of chronic disease in adults was the first disease to be hypothesized by Barker in relation to changes occur during fetal life (52). It is important to refer that premature babies at different gestational ages are metabolically distinct, studied by (49). This study reported associations between the different stages of prematurity and levels of amino acids, enzymes and endocrine markers in a large cohort of infants (49).

The metabolites represent the organism phenotype and when they vary quantitatively or/and qualitatively, they may act as markers of disease or other perturbations (45). In case of PTB metabolic changes can be found in a pre-natal phase (*in utero*), indicating the risk of PTB, or in a postnatal phase (after birth), leading to a better diagnostic and prognostic for the possible consequences of PTB (45). These metabolites, as was previously referred, can

be detected in biofluids and the kind of biological sample depends on the objective of the study (33). Biofluids can be amniotic fluid, maternal blood [MB], maternal urine [MU] (collected before birth), umbilical cord blood [UCB], MB, MU and newborn urine [NU] (collected at or after birth) (42).

In paediatrics and neonatology, urine is especially important to use since it can be collectable non-invasively, it is abundant, rich in metabolites (24) and its composition reflects many of the metabolic processes in the newborn. Therefore, it may provide information about the overall metabolic status or about development and metabolic changes occur in case of disease (38,45,53). However, urine analysis can become difficult to interpret because of the chemical concentration of their components. So, it is important to consider and uniformize several factors such as sample collection, data acquisition and processing parameters (24). Although urine is the biofluid indicated for metabolomic studies in PTB, is in the others biofluids have been used in this area (37,42).

In a recent study a typical metabolic profile of the urine of a term newborn (>37 g.w.) serves as control when compared with preterm newborns (*Figure 5*) (54).

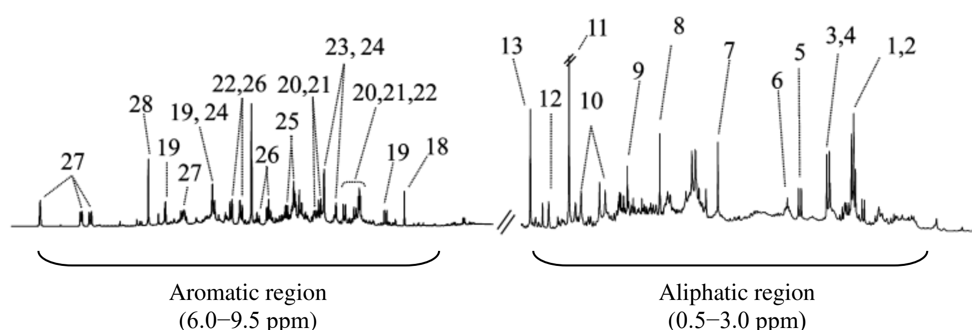


Figure 5 **-** Average 500 MHz ^{1}H NMR aliphatic and aromatic regions of urine of controls newborns. Some assignments are noted: 1:3-aminoisobutyrate, 2:3-hydroxybutyrate (3-HBA), 3: lactate, 4: threonine, 5: alanine, 6: adipate, 7: acetate, 8: acetone, 9: succinate, 10: citrate, 11: dimethylamine (DMA), 12: methylguanidine, 13: dimethylglycine (DMG), 14: taurine, 15: creatine, 16: glucose, 17: lactose, 18: fumarate, 19: N-methyl-2-pyridone-5carboxamide (2-Py), 20:4-hydroxyphenylacetate (4-HPA), 21: tyrosine, 22:4-hydroxyhippurate, 23:1-methylhistidine, 24: histidine, 25: phenylacetylglutamine (PAG), 26: hippurate, 27: N-methyl-nicotinamide (NMND), 28: formate [Adapted from (54)].

Another report (50) studied the changes occurring in the metabolic profile of newborns urine in the first weeks of life. The study considered newborns (n=10) born at term (g.w. not defined) and the samples were analysed at 0, 14 days, 8 weeks and 14 months postpartum. Newborns metabolism changed rapidly in the first few weeks of life due to fast adaptation of physiology. At the time of birth, newborn metabolic profiles are very similar with that of the mother (urine sample of the mother collected at the same time of the newborns sample). This similarity can be explained by the recent dependence of placenta nutrients. After 14 days a lower intensity of metabolites is visible, specially creatinine and hippurate. The aromatic region is reduced and increased betaine was found. At 14 months the urine metabolic profile showed higher concentration levels in the aromatic region which were caused by food-driven gut microorganisms.

1.4.2. Application of biofluid metabolomics to premature newborns

There are many studies that have analyzed and reported that physiological variables can influence the metabolic profiles in addition to pathologic conditions (38). These physiological conditions can be denominated as confounders factors, and include gender, delivery mode, gestational age at birth, day-of-life at collection and the feeding given by mother (54,55). These factors may influence urine composition and hence the biomarkers for long-term diseases associated to PTB.

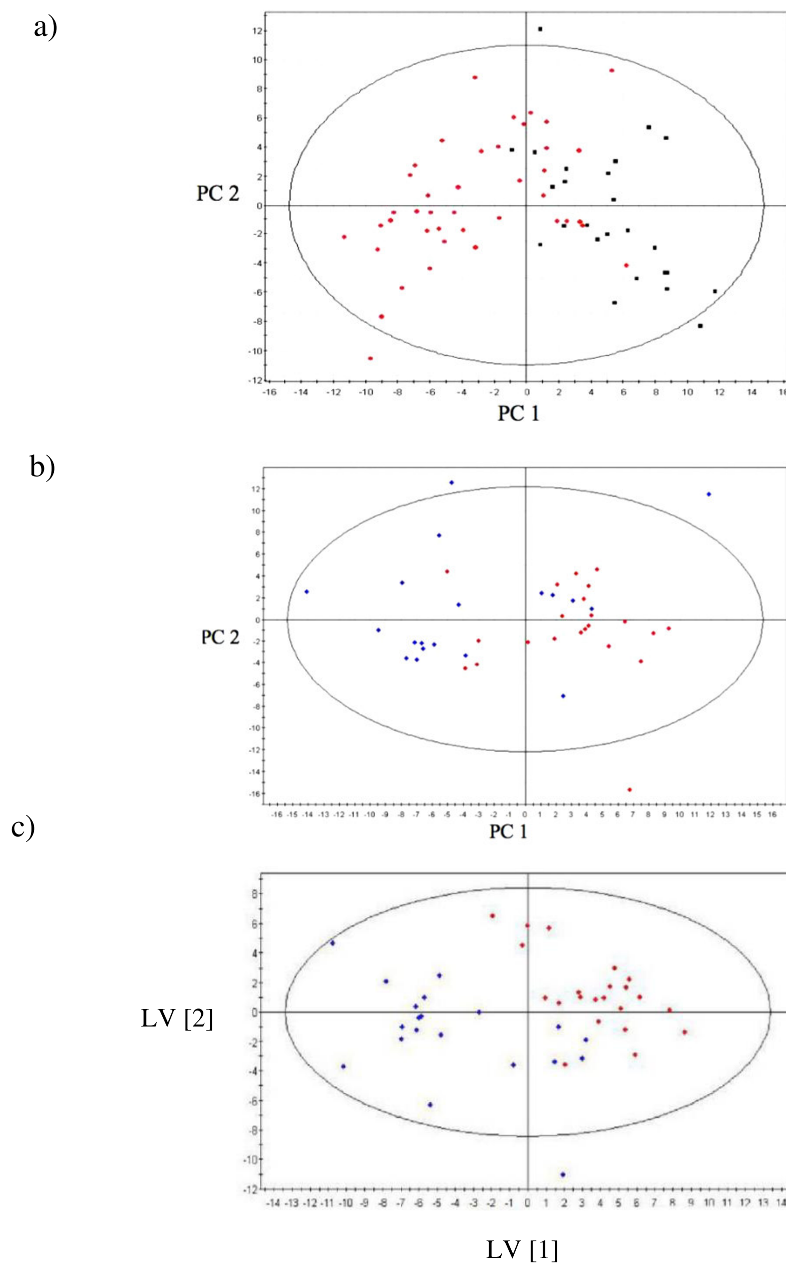
So far, a metabolic profile of the PTB newborn was studied at the time of birth through the analysis of amniotic fluid (AF) (using MS spectrometry), umbilical cord blood (UCB) and maternal blood (MB) (using both NMR spectroscopy and MS spectrometry) in order to find a metabolic signature of PTB at birth and a metabolic relationship with the mother.

Postnatal metabolism associated to PTB newborn has been studied through the analysis of newborn urine (NU) and newborn blood (NB) in the first days after birth. The first studies have been focused on ^{1}H-NMR analysis of NU carried out in preterm infants (53,56,57), followed by a few studies that used ^{1}H-NMR metabolomic analysis of NU in preterm infants (54,43). In relation to studies in NB the analysis is performed in blood spots and used MS analysis, sometimes in tandem with other targeted techniques, such as enzymatic kits, and used usually in very large cohorts. This set of reports on metabolomics in NB established that the capacity for an accurate gestational age estimation at birth (58–62). The principal variations occur in arginine, leucine, valine, glutamine, pipecolic acid, proline, homocysteine, carnitines and hormones.

In the following sections the results found in recent reports are described postnatal metabolomics studies in urine using ^{1}H NMR spectroscopy in preterm babies (Subchapter 4.3.1.), as well as some reports on neonatal diseases (Subchapter 4.3.2.).

1.4.2.1. Premature compared to healthy newborns

The first studies reported on (53, 56, 57) have small cohorts (n=8, 7 and 5 preterms, respectly) while the newer ones have used large cohorts (n=17 preterms to 46 controls) (54), which in newborns studies is a satisfactory cohort. Metabolic changes were found in taurine (↑), trimethylamine *N*-oxide (TMAO) (↑) and *myo*-inositol (↑) (57); *N*-methyl-nicotinamide (NMND) (↑) and taurine (↑) (53). The urine metabolic profiles in term (n=26) and PTB infants (n=41) was studied in order to reveal distinct gestational age classes (38). *Figure 6* a) represents the PCA scores plot of term and preterm newborns where it is possible to observe two clusters with overlapping regions. Then two PTB groups were divided: Group A n=22 (23 to 32 g.w.); Group B n=19 (33 to 36 g.w.). PCA and a PLS-DA scores plots (*Figure 6* b) and c) respectively), revealed two distinct clusters with an overlapping region.



(R^2X: 0.2; R^2Y: 0.74; Q^2:0.21)

Figure 6 – a) PCA scores plot from the full dataset. (▪ term infants, ▪ PTB infants); b) PCA scores plot from the PTB samples. (▪ PTB infants of 23 to 32 g.w. (group A), ▪ PTB infants of 33 to 36 g.w. (group B); c) PLS-DA scores plot from the PTB samples. (▪ PTB infants of 23 to 32 g.w. (group A), ▪ PTB infants of 33 to 36 g.w. (group B). [Adapted from (38)].

Metabolite alterations occur in hippurate, triptophan, phenylalanine, malate, tyrosine, hydroxibutyrate, *N*-acetylglutamate and proline (the variations are not described) (38).

One last study investigated the impact of prematurity on newborns urine composition and in *Figure 7* a) a PLS-DA scores plot of controls (term newborns) (n=46) compared with premature newborns (n=17) shows a separation of two groups. Group separation is due to a signature that was found in the PTB newborn urine comprising variations with high contribute levels of 3-aminoisobutyrate, 3-hydroxybutyrate (3-HBA), *cis*-aconitate, ethanolamine, *N*-methyl-2pyridone-5-carboxamide (2-Py) and NMND along with some unassigneds and several aliphatic and aromatic regions. These metabolites (with high VIP values) can be seen in *Figure 7* b) where is represented a variable importance to projection (VIP) wheel of the NMR metabolite signatures obtained for prematurity. In the inner circle are the average ^{1}H NMR controls spectrum with the respective ppm scaling.

Notably, there are two regions in the *Figure 7* that are a slightly separated from the corresponding group showing that gestational age or birth weight are not the only variables that could affect these metabolic differences. These differences can be possibly cause by other confounding effects. Despite this, this study provides a metabolic signature much more complete compared with the other studies (38,53), making it possible to be used in the PTB monitoring until theoretical term time (37).

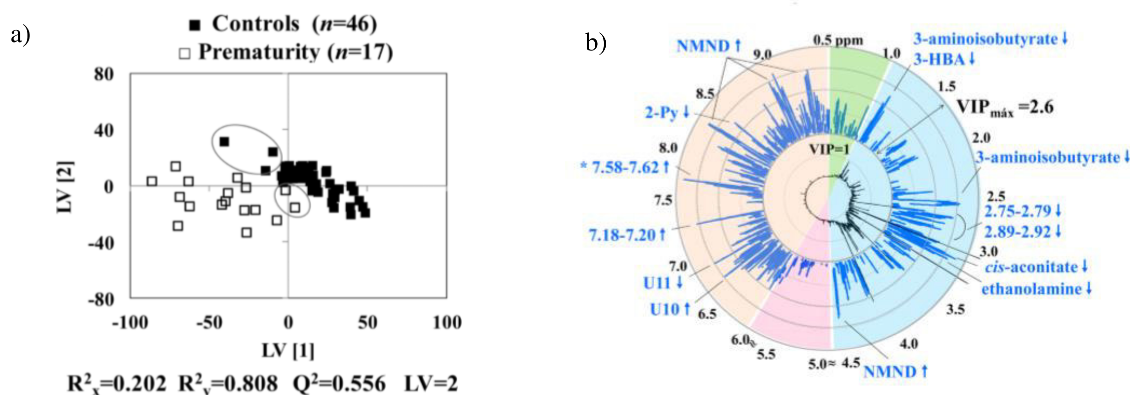


Figure 7– a) PLS-DA scores plot obtained for prematurity (□, n=17) compared with controls (■, n=46); b) VIP-wheel representation of the NMR metabolite signatures obtained. *Indicates a region with high VIP value but no statistical relevance (Pvalue>0.05) [Adapted from (54)].

1.4.2.2. Premature compared to newborns affected by other disorders

Babies that are born preterm, may require extra care support for complications (such as respiratory depression, NEC, perinatal asphyxia, jaundice, acute kidney diseases (AKDs),

neonatal sepsis, among others. that may develop. These health complications can appear also in newborns at term as general stress effects. Furthermore, there are some diseases that are congenital disorders of metabolism such as, inborn errors of metabolism (disease caused by a deficient enzyme activity in intermediate metabolism (63) and small or large for gestational age (SGA/LGA)). So, it is important compare these diseases with prematurity to find a specific signature to PTB. There is a lack of literature that used urine in the ^{1}H-NMR based metabolic analysis of newborns, however the main studies and findings are listed in *Table 2*.

Table 2 – Postnatal metabolomics studies in urine using ^{1}H-NMR spectroscopy. Abbreviations: NMND- *N*-methyl-nicotinamide; 2-PY- *N*-methyl-2-pyridone-5-carboxamide; 3-HBA- 3-hydroxybutyrate; 4-DTA- 4-deoxy-threonic acid; DMG- dimethylglycine; TMAO- trimethylamine *N*-oxide; DMA- dimethylamine; LGA- low for gestational age; BDP- Bronchopulmonary Dysplasia.

Disease	**Patients**	**Sample**	**Technique**	**Metabolites**	**Reference**
Prematurity	n=17 preterm n=46 controls	urine	^{1}H-NMR	Increased: NMND, Decreased: 3-aminoisobutyrate, 3-HBA, *cis*-aconitate, ethanolamine, 2-Py	(54)
LGA	n=18 newborns LGA n=46 controls	urine	^{1}H-NMR	Increased: 2-Py, DMG an hippurate; Decreased: *myo*-inositol	(54)
Perinatal asphyxia	n=6 newborns n=8 controls	urine	^{1}H-NMR	Increased: threonine, 3-hydroxyisovalerate, lactate, glucose, TMAO; Decreased: acetate, succinate, DMA, citrate, DMG, creatine, creatinine, betaine, *cis*-aconitate, *trans*-aconitate, urea, formate	(64)
BDP	n=18 BDP newborns n=18 controls	urine	^{1}H-NMR	Increased: lactate, taurine, TMAO, *myo*-inositol; Decreased: citrate, glucanate	(65)
Respiratory depression	n=10 newborns RD n= 46 controls	urine	^{1}H-NMR	Increased: glycine, taurine, *myo*-inositol Decreased: 4-DTA and threonine	(54)

According to the studies relating to NU by ^{1}NMR spectroscopy, is possible to observe that metabolic changes for the indicated diseases are distinct from those found for prematurity.

1.5. Aims of thesis

This work has entailed two goals:

- The first was to establish a metabolic signature of PTB NB urine on the first days of life. This information would provide the study of metabolic impact of prematurity on babies at different gestational ages.
- The second was to evaluate the metabolic impact of prematurity on preterm NBs over time. This study allows a monitoring PTB NBs from birth until term time, in order to obtain a urinary metabolic trajectory that allows to be found new biomarkers of deviant behaviors indicative of health complications.

Chapter 2. Experimental Section

This work was made in collaboration with the Maternity Bissaya Barreto (MBB), University Hospital Center of Coimbra (CHUC), were NB urine samples were collected some in 2016 and others in 2017, under CHUC ethical committee approval (refs 18/04 and 29/09), with parental informed consents obtained for each infant. Clinical information of PTB and NBs was obtained from obstetrical and neonatal medical records and individual questionnaires filled at the time of collection.

2.1. Sampling

2.1.1. Sample and metadata collection

Newborn urine was collected using the cotton-ball method that consists in placing of a sterile cotton-ball inside the newborns diaper, for up to 3 h, to retain urine (approximately 1.5 mL). Urine was transferred to a sterile flask and frozen at -20°C for up to 24 h and then transferred to -80°C where they were stored until analysis. Samples were transported to CICECO-Chemistry Department, at University of Aveiro (UA), where they were stored (-80ºC) and analysed. Each sample was given a code that corresponds to each NB process number to ensure patient confidentiality. Clinical information on mothers prenatal and pregnancy history, as well as NB clinical metadata at the time of birth were provided by the clinical professionals at MBB.

2.1.2. Definition of sample groups

This thesis is divided into two groups: Study 1 which aimed to define a metabolic signature of preterm birth (PTB) in NB first urine collection at birth time and Study 2 which aimed to establish a urinary metabolic trajectory of PTB NB during hospital stay, to identify putative markers of deviant behaviours indicative of health complications during that period. For both studies a control group was used, defined as healthy NBs born at term time with no pregnancy disorders detected.

Study 1:

In *Table 3* the first NB urine samples collected for a term group (37-40 g.w.) and a PTB group (26-36 g.w.) are listed, as well as gender, type of delivery, corresponding day of life at sampling, birth weight, maternal age, gravidity (no. pregnancies) and parity (no. deliveries) of each subject. In this study there were 3 PTB NBs born of mothers with arterial

hypertension, 5 with gestational diabetes, 3 with pre-eclampsia, 1 with periventricular haemorrhage (HPV), 1 with previous abortions history and 2 with previous membranes rupture. In relation to PTB NBs at the time of birth there were 7 PTB babies with transition tachypneia, 8 with hyperbilurrubinemia, 2 with hyponatremia, 1 with fetal growth restriction, 2 with infection, 1 with apnea, 1 with pneumothorax and 1 with surfactant deficit. Noting that in this study there are 10 twin PTB NBs were 4 pairs are twin brothers and other 2 corresponds to different brothers.

Table 3- List of first NB urine samples collected for control group and PTB group used for Study 1. EPT- extremely PTB; VPT - very PTB; MLPT - Moderate to late PTB; M - male; F - female; VD - vaginal delivery; CS - cesarean section. Median valeus are shown between brackets.

Group	Control	Prematurity
n	46	36 (2 EPT; 6 VPT; 28 MLPT)
M/F	23/23	22/14
VD/CS	29/17	16/18
Day of life	1-4 (2)	1-6 (2.6)
Gestational age (g.w.)	37-40 (39)	26-36 (34)
Birth weight (kg)	2.46-3.69 (3.11)	0.79-3.28 (2.15)
Maternal Age (years)	16/43 (31)	23-41 (33)
Gravidity/Parity	1-6 (2)/0-3 (1)	1-6(2)/0-2 (1)

Study 2:

In *Table 4* the all NB urine samples collected for a PTB group (26-35 g.w.) are listed, as well as gender, type of delivery, corresponding day of life at sampling, birth weight, maternal age, gravidity (no. pregnancies) and parity (no. deliveries), hospitalization time and last day of urine collection.

Table 4 - List of all NB urine samples collected as function of time for PTB group used for Study 2. NB - Newborn; n –number of samples collected; M - male; F - female; VD - vaginal delivery; CS - cesarean section; EPT – Extremely preterm; VPT – Very preterm; MLPT – Moderate to Late preterm.

PTB	n	Days of collection	M/F	VD/CS	Gestational age (g.a.)	Birth weight (kg)	Maternal Age (years)	Gravi-dity/ Parity	Hospi-talization	Last collection Day
1	2	1,5	M	VD	27	1.520	40	3/2	8	5
2	3	3,4,5	M	VD	35	2.980	32	1/1	9	5
3	3	1,5,6	M	CS	33	2.190	30	2/1	18	6
4	2	1,6	F	CS	31	1.620	26	1/0	8	6
5	9	2,3,4,5,9,17 ,19,23,24	F	VD	26	0.790	39	6/0	55	25
6	2	4,5	F	VD	35	2.406	35	2/2	9	5
7	3	1.9.10	F	CS	34	1.850	29	3/2	15	10
8	2	1,9	M	CS	34	2.345	29	3/2	15	9
9	3	1,6,13	M	VD	31	1.860	38	2/2	28	13
10	3	1,6,14	M	VD	31	1.495	38	2/2	28	14

For this study the PTB group contain two EPT (PTB 1 and 5), three VPT (PTB 4, 9 and 10) and five MLPT (PTB 2,3,6,7 and 8) and a total of 30 samples.

2.2. NMR spectroscopy

2.2.1. Sample preparation for NMR analysis

Before analysis, samples were thawed at room temperature (±20-30 min) and 600 µL were centrifuged (8000 rpm, 5 min). The pH was measured before and after the centrifugation to see the pH of supernatant. Then 540 µL of supernatant were added to 60µL of a 1.5 M KH_2PO_4/D_2O buffer pH 7.0, 0.1% Na^+/3-trimethylsilyl-propionate (TSP) followed by pH readjustment to 7.00 ± 0.02 with KOD (4 M) and/or DCl (4 M). The mixture was centrifuged (8000 rpm, 5 min) and 550µL of the supernatant were transferred to a 5 mm NMR tube.

2.2.2. Acquisition and processing of NMR data

NMR spectra were recorded on a Bruker Avance III HD 500 spectrometer, operating at a proton frequency of 500MHz, and acquired at 300 K. Spectra were then manually phased and baseline corrected, and chemical shifts calibrated internally to TSP at $\delta = 0.0$ ppm on Topsin. The regions of water (4.68-5.05 ppm) and urea (5.66-5.96 ppm) were excluded to minimize spectra variations due to water suppression and different urea levels, respectively. The acquisition parameters used can be found in reference 54.

2.2.3. Pre-processing and multivariate analysis of NMR data

Proton 1D NMR spectra were aligned using a recursive segment-wise peak alignment on AMIX, after removal of water and urea regions, and normalized by Probabilistic quotient normalization (PQN) using MATLAB software.

MVA was performed through SIMCA-P software, by applying PCA and PLS-DA to the matrix previously aligned and normalized. Model robustness was evaluated in terms of Q^2. The corresponding loading plots were back-transformed by multiplying the loading weight [w] by the variable standard deviation. The relevant peaks found in the loadings plots were integrated in the original spectra using the AMIX program. Integrals were carried out using Python software for determine the p-value in order to reveal the statistical significance ($p\text{-value}<0{,}05$). Then, for each set of metabolite, the biological relevance was determined by the effect size, calculated according to equations of the article (66).

Peak assignments were carried out with basis on literature, 2D NMR experiments (TOCSY, HSQC), consultation and comparation with databases (Bruker Biorefcode database and HMDB – human metabolome database)

Chapter 3. Typical NMR spectra of newborn urine and assignments

This chapter presents the results obtained for the characterization of healthy newborns urine composition at term time (>37 g.w.), through 1D and 2D NMR experiments and the resulting list of metabolites identified.

Urine is a complex biofluid and its typical ^{1}H NMR spectrum contains hundreds of peaks arising from numerous metabolites. A typical ^{1}H NMR urine spectrum of newborns at term time is shown in *Figure 8* along with some assignments. It is clearly visible that major peaks occur in aliphatic (0,5-3 ppm) and sugar regions (3-5 ppm), while in the aromatic region (6,5-9,5 ppm) signals are weaker.

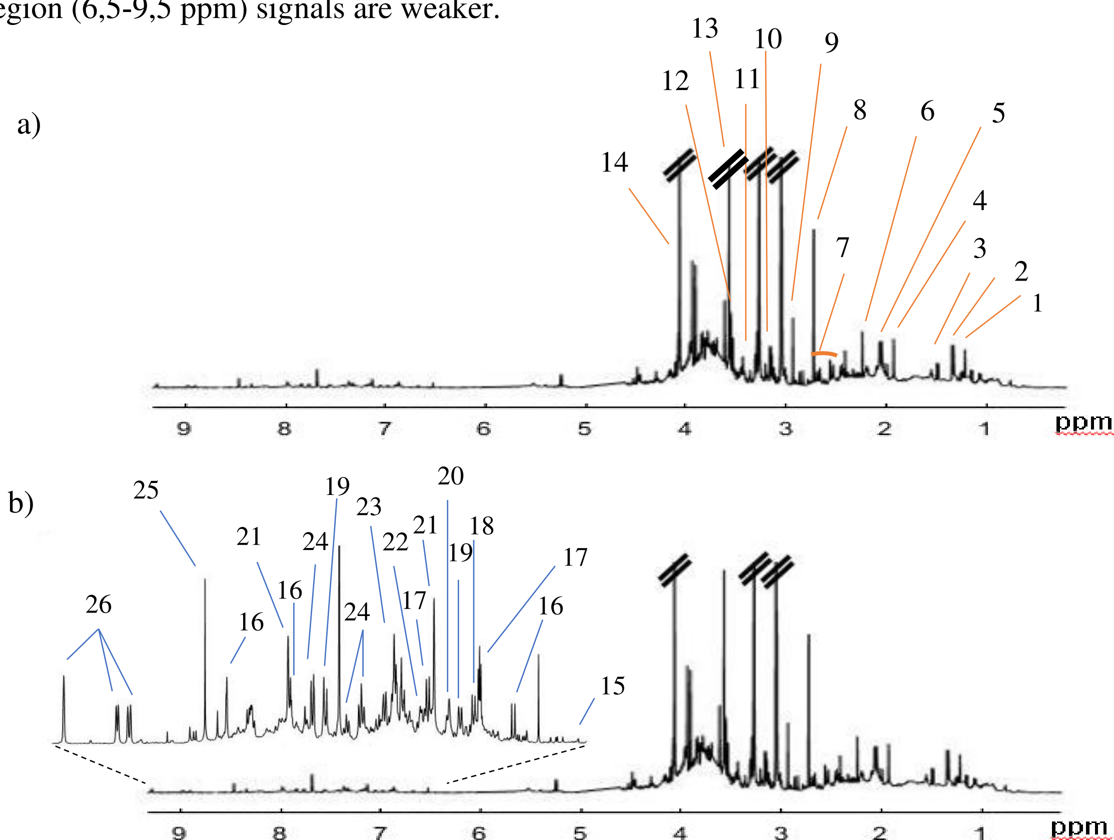


Figure 8 – a) Average 500 MHz ^{1}H NMR spectrum of term newborns (control group), with some peak assignments. Legend: 1- 3-HBA, 2- lactate, 3- alanine, 4- acetate, 5- *N*-acethyl neuraminic acid, 6- acetone, 7- citrate, 8- DMA, 9- DMG, 10- ethanolamine, 11-taurine, 12- m-inositol, 13-glicine, 14- creatinine; b) Expansion of region 6,5-9,3ppm with some peak assignments. Legend: 15- fumarate, 16- 2-PY, 17-4-HPA, 18-tyrosine, 19- 4-OH-hippurate, 20- 1-methyl-histidine, 21- histidine, 22- tyrosine, 23- PAG, 24- hippurate, 25- formate, 26- NMND.

The complexity and extensive peak overlap are clearly visible so, 2D NMR is extremely important to identify some unknown metabolites. In *Figure 9 and 10* expansions of 2D NMR spectra TOCSY and HSQC are shown respectively, with indication of some assignments.

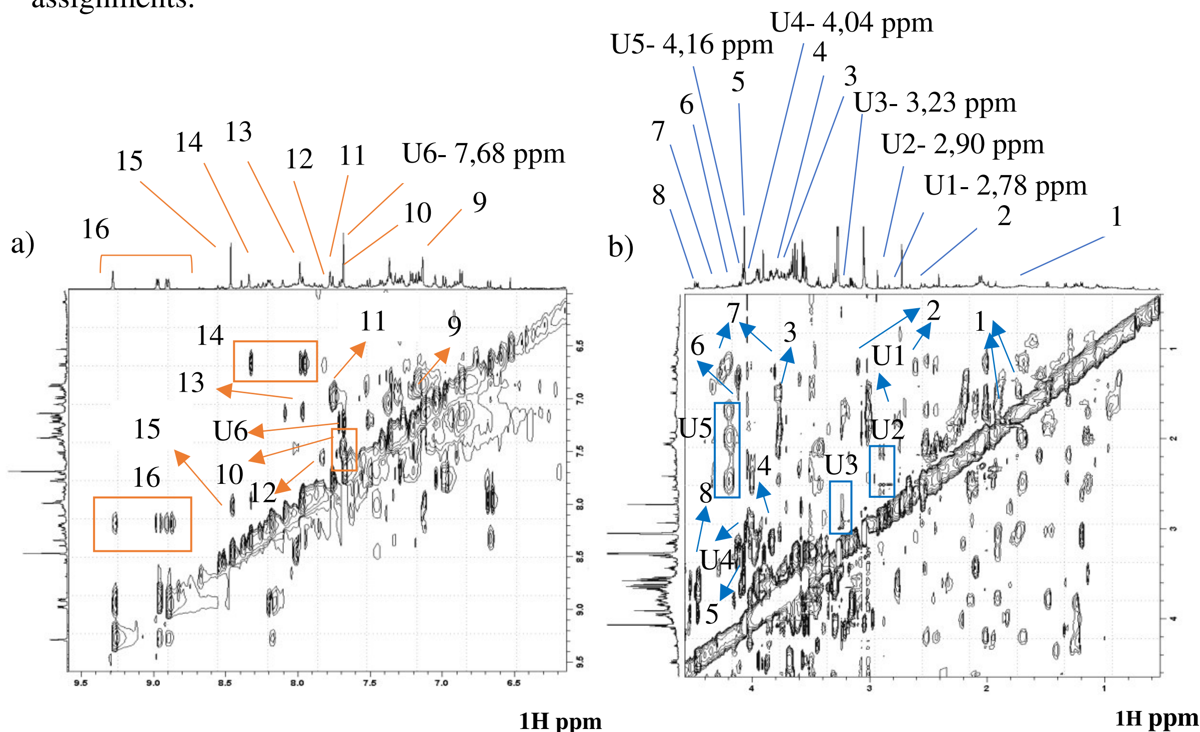


Figure 9 – Expansions of the 2D NMR TOCSY spectra of urine in the regions from a) 5,1-9,6ppm and b) 0,6-4,6ppm. Legend: 1-lysine; 2- 3-aminoisobutyrate; U1- 2.78ppm (possible correlation with *N*-acethylneuraminic acid at 2.05ppm); U2- 2.90ppm (possible correlation with 2.56; 2.19; 3.20 (coline)); U3- 3.23ppm (possible correlation with 1.70, 1.92, 2.68, 2.90ppm); 3- alanine; 4- ethanolamine; 5- *m*-inositol; U4- 4.04ppm (possible correlation with 3.05ppm (creatinine)); U5-4.16ppm; (possible correlation with 2.47, 2.00, 1.70ppm); 6- lactate; 7- threonine; 8- Lactose; 9- 4-HPA; 10- IS; U6- 7.68ppm (possible correlation with 7.22ppm); 11- 4-OH-hippurate; 12- hippurate; 13- histidinde; 14- 2-Py; 15- formate; 16- NMND.

1D spectrum shows peak overlap with other ressonances, but with TOCSY the assignment can be facilated, specially when peaks intensity are lower.

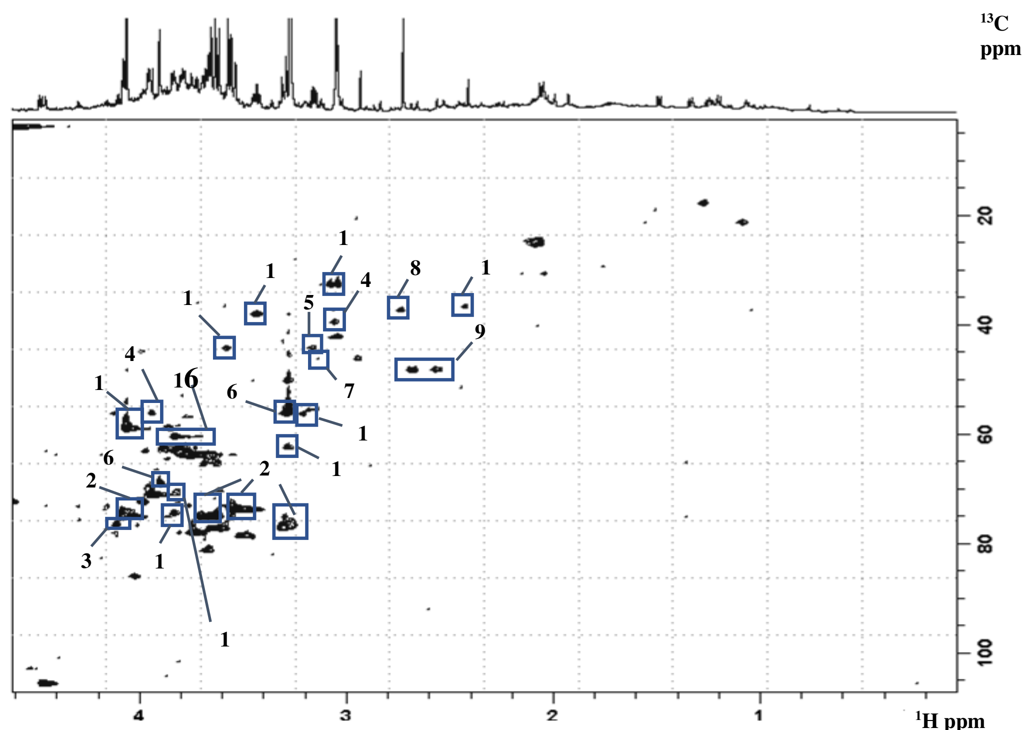


Figure 10 – Expansion of the 2D NMR HSQC of urine in the region from 0,0-4,6ppm. Legend: 1- Creatinine- 3,05 (s)/32,6; 4,06 (s)/59.0; 2- *m*-inositol- 3,29 (t)/76,8; 3,63(t)/74,8; 3,55 (dd)/73,6; 4,07 (t)/74,8; 3- Lactate- 4,11 (q)/76,6; 4- Creatine- 3,04 (s)/39,6; 3,93 (s)/56,5; 5- Ethanolamine- 3,15 (t)/44.0; 3,84 (t)/60,1; 6- Betaine- 3,26 (s)/55,8; 3,90 (s)/69,0; 7- Cis- aconitate- 3,12 (d)/46,2; 8- DMA- 2.72ppm (s)/37.4; 9- Citrate- 2,54 (d)/48,2; 2,67 (d)/48,2; 10- Succinate- 2.41ppm (s)/36.2; 11- Glycine- 3,57 (s)/44,5; 12- Choline- 3,20 (s)/56,5; 13- TMAO- 3,27 (s)/62.4; 14- DMG- 2,93 (s)/46,2; 3,72 (s)/63,6; 15- 4-DTA- 3,84 (d)/73,9; 16- Glutamine- 3,79 (t)/60,7; 17- Glucose- 3,43 (m)/38,4.

In total, 49 metabolites were identified in NB urine, including amino acids and derivatives, organic acids and sugars, all having been previously identified in NBs urine (54). Peak identification was made by comparison with databases, such as Human Metabolome database (HMDB), Bruker Biorefcode and previous report (54). In *Table 5* the complete list of assignments is shown along with the corresponding ^{1}H chemical shifts, multiplicity, assignment and, when possibly, ^{13}C chemical shifts.

Table 5 – Main peak assignments in the ^1H NMR spectrum (500 MHz) of newborn urine, at pH 7.00±0.02. (s: singlet, d: doublet, t: triplet, q: quartet, dd: doublet of doublets, m: multiplet).

Metabolite	**$\delta_{H\ ppm}$ (multiplicity)**
1-methylhistidine	3.16 (dd, β'CH_2); 3.72 (s, CH_3); 3.97 (dd, αCH_2); 7.04 (s, ring); 7.78 (s, ring)
2-hydroxyisobutyrate (2-HIBA)	1.36 (s, CH_3)
3-aminoisobutyrate	1,19 (d); 2,61 (m); 3,09 (dd)
3-hydroxisovalerate (3-HIVA)	1,27 (s); 2,37 (s, αCH_2)
3-hydroxybutyrate (3-HBA)	1.20 (d, CH_3); 2.31 (m, CH_2); 2.41 (m,CH_2); 4.15 (m,CH)
4-deoxythreonicacid (4-DTA)	1.23 (d, γCH_3)
4- hydroxyhippurate	3.96 (s, CH_2); 6.98 (d, C3H, C5H ring); 7.76 (d, C4, 2H, C6H ring)
4-Hydroxyphenylacetate (4-HPA)	6.86 (d, C3H, C5H ring); 7.16 (d, C2H, C6H ring)
Acetate	1.92 (s, βCH_3)
Acetone	2.23 (s, CH_3)
Adipate	1.55 (m, βCH_2); 1.91 (m, αCH_2)
Alanine	1.49 (d), 3,78 (q)
Ascorbate	3.76 (m, CH2(OH)); 4.01 (m, CH(OH)); 4.52 (d, C1H)
Betaine	3.26 (s, CH_3); 3.90 (s, CH_2)
Choline	3.20 (s, $N(CH_3)_3$)
Cis-aconitate	3.12 (d, CH)
Citrate	2.54 (d, a,b CH); 2.67 (d, a',b' CH)
Creatine	3.04 (s, NCH_3); 3.93 (s, NCH_2)
Creatinine	3.05 (s, NCH_3); 4.06 (s, NCH_2)
Dimethylamine (DMA)	2.72 (s, CH_3)
Dimethylglycine (DMG)	2.93 (s, $(CH_3)_2$); 3.72 (s, CH)
Ethanolamine	3.15 (t); 3.83 (t)
Formate	8.46 (s, CH)
Fumarate	6.52 (s, CH)
Galactose	3.64 (dd, C3H); 3.75 (m, C1H, C2H, CH_2); 3.83 (m, C3H); 3.99 (d, C2H); 4.07 (t, C1H); 4.59 (d, CH_2); 5.28 (d, C5H)
α-Glucose	4.65 (d, C1H)
β-Glucose	3.43 (t, C4H); 3.54 (dd, CH); 5.24 (d, C1H)

(Continues on next page)

(Continuation of the Table 5)

Metabolite	$\delta_{H\ ppm}$ (multiplicity)
Glutamine	2.14 (m, βCH_2); 2.45 (m, γCH_2)
Glycine	3.57 (s, αCH_2)
Hyppurate	3.96 (d, CH_2); 7.55 (t, C4H, C6H ring); 7.64 (t, C3H, C5H ring); 7.83 (d, C4H); 8.55 (s, NH)
Histidine	7.13 (s, C4H ring); 7.98 (s, C2H ring)
Indoxyl Sulfate (IS)	7.36 (s, C2H); 7.50 (d, C6H); 7.70 (d, C9H)
Lactate	1.33 (d, CH_3); 4.11 (q, CH)
Lactose	3.55 (m, C'2H); 3.59 (m, C2H); 3.66 (m, C'3H, C3H, C5H); 3.73 (m, C'6H, C'5H); 3.79 (m, C6H); 3.86 (m, C6H, C3H); 3.94 (m, C6H, C'4H, C4H); 4.46 (d, C'1H); 5.25 (d, C1H)
Lysine	1.48 (m, γCH_2); 1.73 (m, δCH_2); 1.92 (m, βCH_2)
Methylguanidine	2.83 (s, CH_3)
Myo-inositol	3.29 (t, C5H); 3.55 (dd, C1H, C3H); 3.63 (t, C4H, C6H); 4.07 (t, C2H)
N-methyl-2-pyridone-5-carboxamide (2-Py)	6.67 (d, C3H ring); 7.97 (dd, C4H ring); 8.33 (d, C6H ring)
N-methylnicotinamide (NMND)	4.48 (s, NCH_3); 8.90 (d, C4H ring); 8.97 (d, C6H ring); 9.27 (s, C2H ring)
Phenilacetylglutamine (PAG)	1.93 (m, βCH_2); 4.18 (m, αCH); 7.36 (m, C2H, C4H, C6H ring); 7.43 (m, C3H, C5H ring)
Sucrose	5,41 (d)
Scyllo-inositol	3.36 (s, CH)
Succinate	2.41 (s, CH_2)
Taurine	3.26 (t, CH_2SO_3); 3.43 (t, NCH_2)
Threonine	1.33 (d, CH_3); 3.60 (d, βCH); 4.27 (dd, αCH);
Trigonelline	4.43 (s, CH_3); 8.84 (br, C2H, C4H ring); 9.12 (s, C6H ring)
Trimethylamine-*N*-oxide (TMAO)	3.26 (s, CH_3)
Tyrosine	6.90 (d, C3H, C5H ring); 7.19 (d, C2H, C6H ring)
Xylose	3.43 (t, C4H); 3.54 (dd, C3H); 3.63 (m, C6H, C5H, C4H); 3.93 (dd, C6H); 4.59 (d, C2H); 5.21 (d, C2H)

Chapter 4. Cross sectional study of newborns with different development degrees at birth – Study 1

In this chapter the results obtained for the NMR metabolomic study of PTB NBs in order to evaluate the metabolic effect of different stages of prematurity on urine composition are described (*Study 1*). The following section presents the comparation of term NBs and PTB NBs. In order to confirm these changes and reveal a metabolic signature for PTB NBs, PCA were applied with a colour distintion of g.a. in PTB group. Furthermore, a bubble g.a. size ratio was made to observe the trend to approximate/disperse from controls. The next section presents the results obtained for the comparison between PTB at different gestational ages (EPT <28 g.w; VPT 28-32 g.w; MLPT 33-36 g.w) and term NBs, with some visual peak assignments. In this chapter, all the studies made used the first urine collection of each NB sample in order to avoid possible confounder effects. The overall changes found in PTB urine composition and throughout gestational ages along with other possible effects are presented, with results of the metabolic profiling of urine and suggestions of possible metabolic interpretations.

4.1. Visual comparison of spectra per class

In this subchapter different stages of prematurity (EPT, VPT, MLPT), were compared. The samples used in this study were those of the first collection, that is collected on the day of birth (day 1) or as close as possible to it (days up to day 6). In *Figure 11* b), c), d) the average ^{1}H NMR spectra of MLPT, VPT and EPT, are represented respectively, along with visual alterations comparing to controls (*Figure 11* a)).

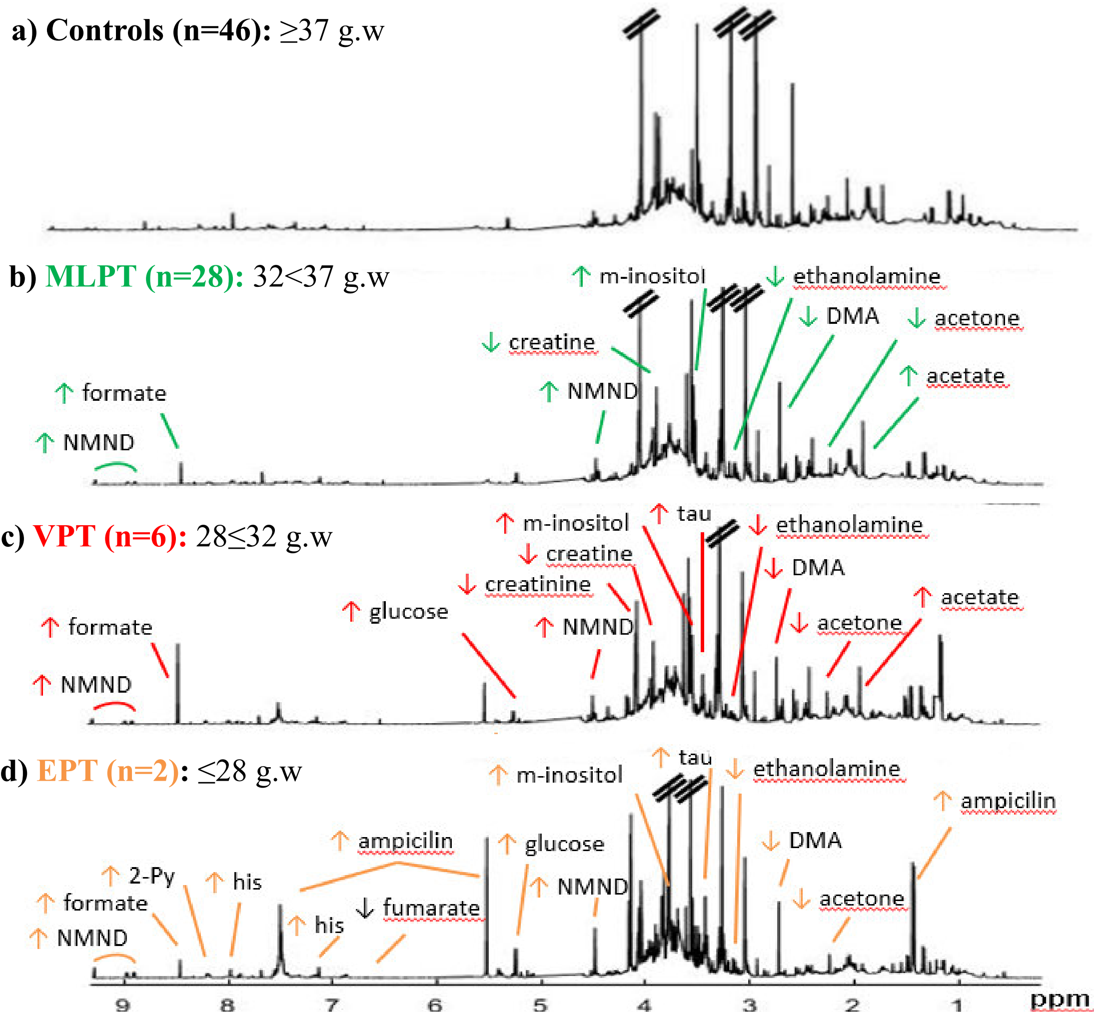


Figure 11 – Average 500 MHz ^{1}H NMR spectra of a) term newborns (control group), and newborns born b) MLPT – moderate to late PTB, c) VPT – very PTB and d) EPT – extremely PTB. Some peak assignments with major visual alterations are indicated in b), c) and d).

Analyzing *Figure 11* it is observed that in EPT case there are more visible peak changes in relation to controls. In relation to increase of NMND, formate and *m*-inositol observed in all stages, the results obtain is in concordance with previous studies using MLPT compared with controls (54,57) and VPT NBs compare to controls (53). Acetone, DMA and ethanolamine were found decreased in all stages, acetate and creatine were found decrease and increased, respectively, on VPT and MLPT. Increase taurine was found in EPT and VPT, which is in concordance with previous studies using MLPT compared with controls (54,57) and VPT NBs compare to controls (53). Citrate was found decreased in EPT but increased

in VPT. The same is visible in succinate which is increased in VPT and MLPT and decrease in EPT. Previous reports (54,57) also found increased citrate and succinate in MLPT compared with controls. Furthermore, in EPT, 2-Py and histidine were found increased and fumarate decreased.

4.2. PCA and PLS-DA analysis of prematurity compare to controls

In order to evaluate the impact of PTB in the urinary metabolome in newborns, confirm and unveil a more complete metabolic picture of premature newborns, scores plots of PCA with all the samples comprising the *Study 1* was carried out PTB (including (EPT (n=2), VPT (n=5), MLPT (n=27)) and NBs at term time (Controls (n=46)) (*Figure 12*). The scores plot of PCA revealed a separation trend between PTB (in general and between the 3 groups of g.a.) and the control group (along PC1) as shown in *Figure 12*. PTB17 (VPT) and PTB26 (MLPT) are clear outliers. To understand why, these two spectra were compared with those of controls and specific changes were found: increased formate, DMG, *m*-inositol, taurine (for both PTB17 and PTB26), galactose, citrate, succinate, acetone (for PTB26), NMND, histidine, 1-methyl-histidine, acetate (for PTB17); and decreased hippurate, creatinine, ethanolamine, DMA, betaine (for both PTB17 and PTB26), 4-HPA (for PTB26) and tyrosine (for PTB17). After inspecting clinical information of these two PTB NBs it was observed that two other samples (PTB5 and PTB43) corresponded to their twin brothers and interestingly, in PCA, those were positioned close to controls. Therefore, these spectra showed evidence of metabolic differences between twin babies.

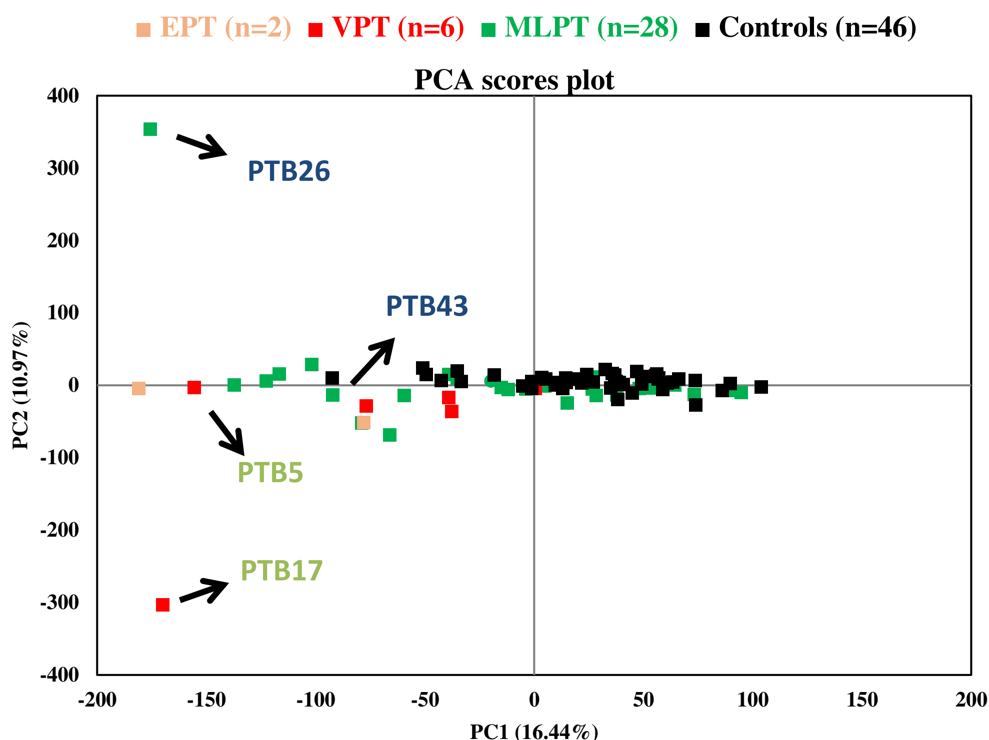


Figure 12 – PCA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ EPT ■ **VPT** ■ **MLPT**) and controls (■).

To understand why there are differences between twin babies, the spectra were compared for the pairs PTB17/PTB5 and PTB26/PTB43 (*Figure 13*). Relatively to PTB17 increased in creatinine, citrate, succinate, acetone; and decreased in NMND, glucose, *m*-inositol. Note that PTB5 has presence of ampicillin and absence of creatine. Relatively of PTB26 metabolic changes was found increased in *m*-inositol, acetate, lactate, propylene glycol and decreased in formate, succinate and acetone.

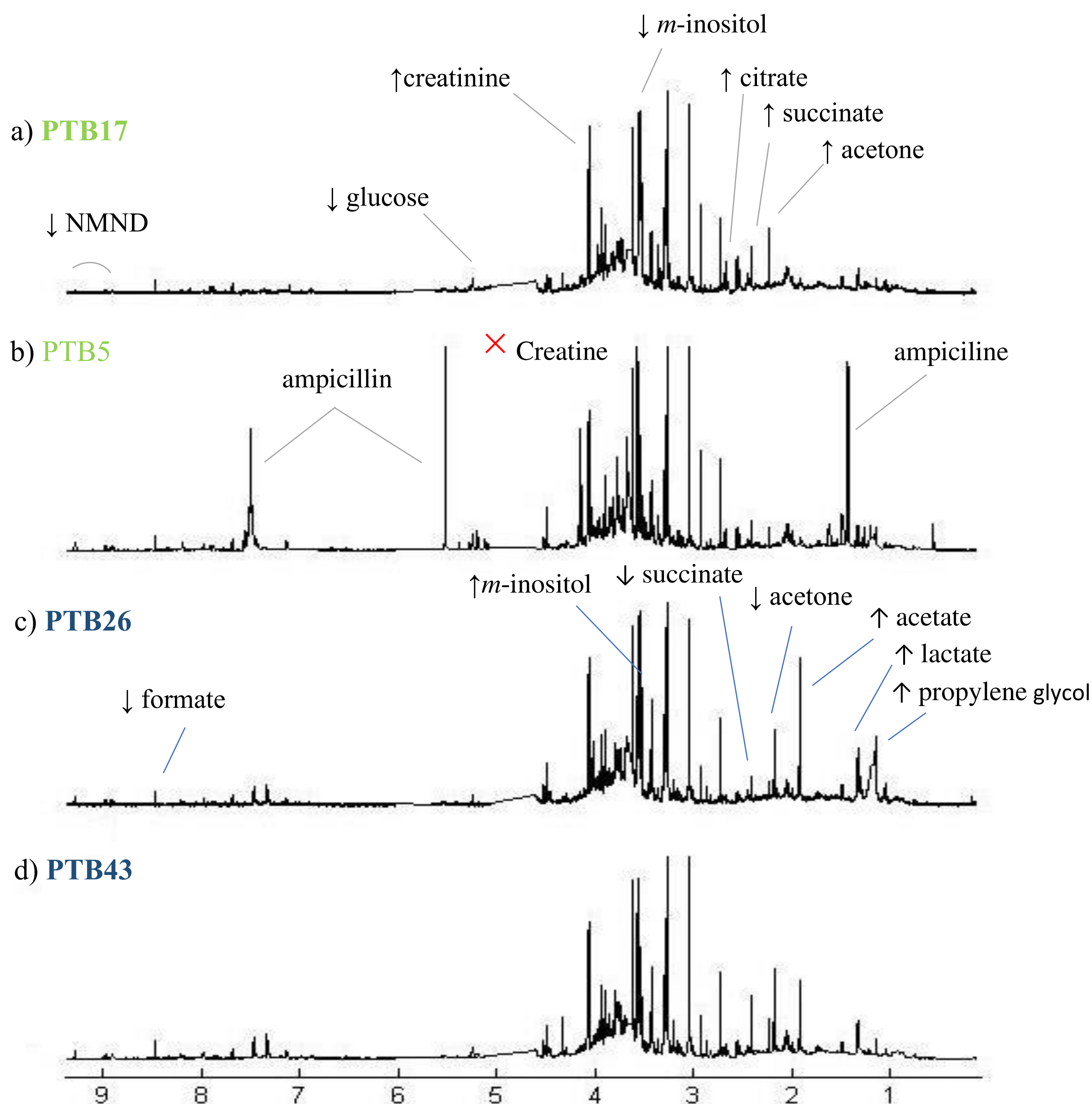


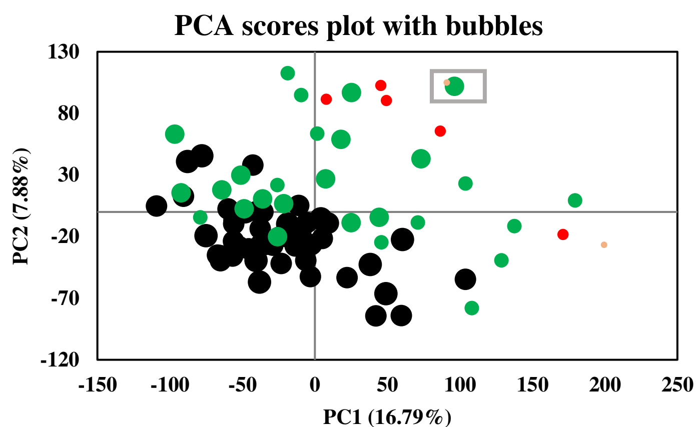
Figure 13 – 500 MHz ^{1}H NMR spectra of a) PTB17, b) PTB5, c) PTB26 and d) PTB43. Differences on peak assignments between PTB17 and PTB5, and between PTB26 and PTB43 are indicated in respective spectra.

After the removal of these two samples considered outliers (PTB17 and PTB26), the separation trend observed was evaluated considering the different gestational ages of PTB group on PCA and PLS-DA scores plots with bubbles where each subject corresponded to a different bubble size reflecting its g.a. (EPT <28 g.w (smaller bublles), VPT 28-32 g.w

(medium bubbles) and MLPT 33-36 g.w (larger bubbles)). *Figure 14* a), b) shows that the plot position of each NB tends to be dependent of g.a., as expected, however some NBs do not follow this trend. In *Figure 14* is also visible (marked with a square) two examples of metabolic differences between stages of prematurity on a significant distant of controls. First between EPT and MLPT (square on *Figure 14* a)), and second with the three stages EPT, VPT and MLPT (square on *Figure 14* b)). These exceptions indicate that the metabolic profile is a more reliable indication of development than clinical staging alone.

● EPT (n=2) ● VPT (n=5) ● MLPT (n=27) ● Controls (n=46)

a)



b)

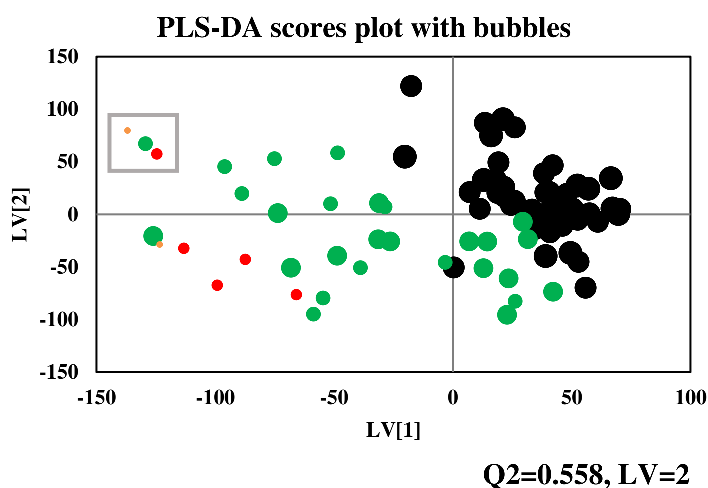


Figure 14 – a) PCA bubbles and b) PLS-DA scores scatter plot bubbles without RN17 and RN26 of ^{1}H NMR spectra of urine from PTB newborns (● EPT ● VPT ● MLPT) and controls (●), where EPT (orange smaller bublle), VPT 28-32 g.w (orange medium bubble) and MLPT 33-36 g.w (orange bubble larger). The squares indicate two examples of metabolic differences between stages.

On PLS-DA is visible a clear separation between classes (PTB and controls), while on PCA some babies MLPT are mixed with the control group.

4.2.1. Proposed metabolic interpretation of changes related with prematurity at birth time

Urinary profiles of twin premature newborns with outlier behaviour

Compared with controls, 2 pairs of twin newborns were found to show significantly different urinary profiles. For each pair, one of the babies had a profile closer to controls (PTB5 on pair PTB17/PTB5 and PTB43 on pair RN26/RN43) and the other baby was significantly deviated from normality (PTB17 on pair PTB17/PTB5 and PTB26 on pair RN26/RN43). Specifically, for the PTB5/PTB17 pair, the deviant baby shown higher excretory levels of creatinine (possibly reflective of altered renal functioned), citrate (involved in the TCA cycle), succinate (involved in the TCA cycle), acetone (ketone body) and lower of NMND (involved in nucleotide metabolism), glucose and m-inositol (precursor of lung surfactant phosphatidylinositol). In addition, for the PTB26/PTB43 pair, the deviant individual showed higher levels of *m*-inositol (precursor of lung surfactant phosphatidylinositol), acetate, lactate, propylene glycol and lower of formate (pyruvate metabolism), succinate and acetone.

Therefore, it is clear that deviations are distinct and not correlated to weight at birth, since that the babies with more weigh are those who are deviated from controls (PTB17 – 1,860 Kg; PTB26 – 2,345 Kg). Also, differences could be due to medication and indeed PTB5 was taking ampicillin which is the neonatal antibiotic most used in INCU (Intensity Care Unity) for infections treatment. This may indicate that this PTB has some infection and these higher values may contribute for alteration of other metabolites such as creatine. However other 3 PTBs are not taking medication, so its use cannot be considered to explain the metabolic differences found. PTB17 were detected with transient tachypnea which is the earliest symptom of respiratory diseases, while PTB5 were detected with SDR (respiratory distress syndrome) which is a respiratory disease caused by a lack of surfactants in lungs. This last condition (as previously explained in chapter 1) can be treated using oxiginoterapy (since there is a decrease in oxygenation) and prevent infections (using antibiotics). *m*-inositol as previously referred is a precursor of lung surfactant phosphatidylinositol. The

high levels of *m*-inositol in PTB5 compared with PTB17 may be explained by more changes in lung surfactant or by an alteration in myelination associated with a possible brain insult caused by the oxygen deprivation. Additionally, PTB43 was detected with hyperbilirubinemia.

Metabolism analysis through urine is not enough for observe directly the state of certain NBs, since several hypotheses (such as gender, delivery mode, birth weight and pathological conditions) (54) were exclude through spectra and clinical conditions analyse. Thus, clinic collaboration is required in accompaniment with metabolomic to improved results interpretation in order to obtain better conclusions.

Urinary profiles of the overall group of premature newborns (upon removal of outliers)

After removal of the PTB17 and PTB26, PLS-DA between PTB and controls shows that exists a g.a. dependence on the distance of each PTB NB to the controls, which can be observed by the different colours that represent the stages of prematurity (green squares (MLPT NBs) tends to approach to controls while red (VPT) and orange squares (EPT) tends to be more distant). However not all NBs follow this trend. This can be proved in PCA and PLS-DA scores plots with bubbles, were besides each group colour there are different bubble sizes, and some of them positioned more closed or more distant to controls. So, the g.a. is not the only factor that contributes for these trends. Besides that, there are some reports (54) which conclude that g.a. is affected by confounding effects (such us delivery mode (CS or VD), gender (M or F), day of life at sampling (1 or 2). Although this study used the first NB collection sample, not all the samples collected correspond to day 1, some having been collected up to day 6. For a better understanding of metabolic changes, some proposals are advanced in the next sections.

Energy metabolism

Citrate and succinate are intermediates of the Citric acid cycle (TCA cycle) and were found increased in PTB NBs. These increases are indicative of altered TCA activity. As a ketone body, acetone is produced from the oxidation of fatty acids in the liver. Acetone and acetoacetate were found increased in PTB NBs.

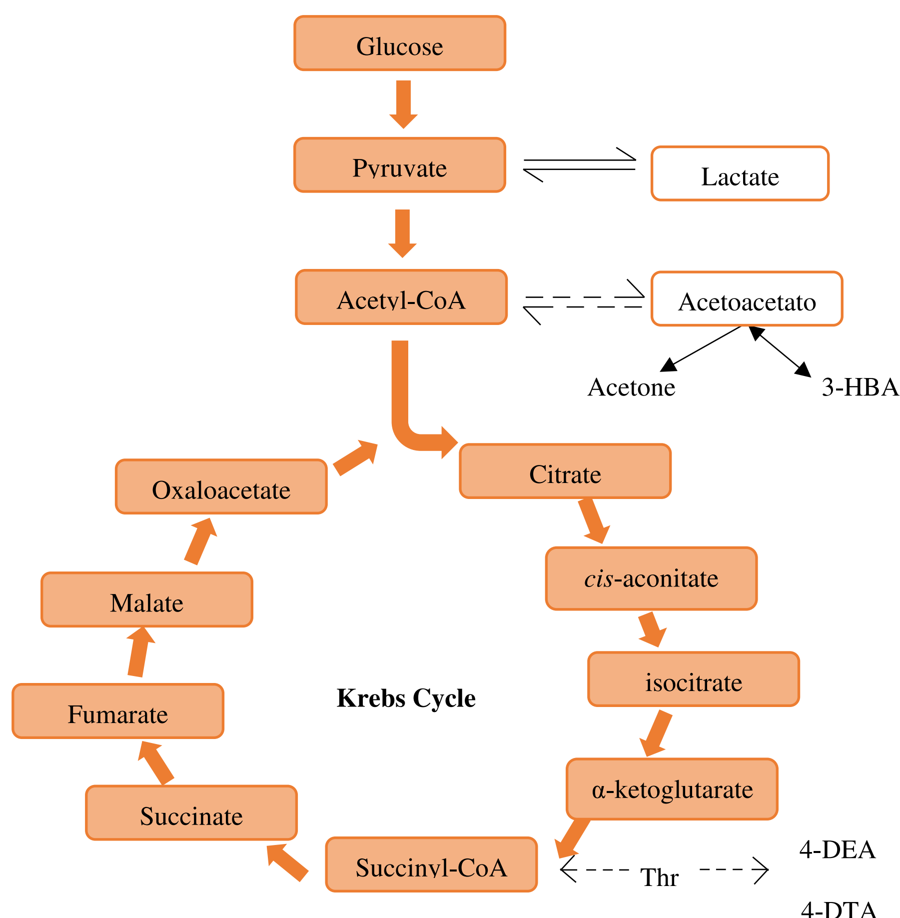


Figure 15 – Schematic representation of the Krebs Cycle and the alterations on the metabolites found in the NMR spectrum of PTB NBs urine. Amino acids letter code: 3-HBA: 3-hydroxybutirate, 4-DEA: 4-deoxyerythronic, 4-DTA:4-deoxythreonic acid

Nucleotide metabolism

NMND is biosynthesed from tryptophan being one of the end products of nicotinamide metabolism along with 2-Py and 4-Py (*Figure 16*). Previously reports found higher NMND in PTB NB (when compared with term NBs) (53,54). In this study, NMND was found increased suggesting that the tryptophan metabolism can be altered, which is in accordance with previous reports (53,54).

Additionally, a report in a study on animals shown that conditions such as diabetes, renal failure and nutrients such as proteins and fatty acids can altered the tryptophan pathway.

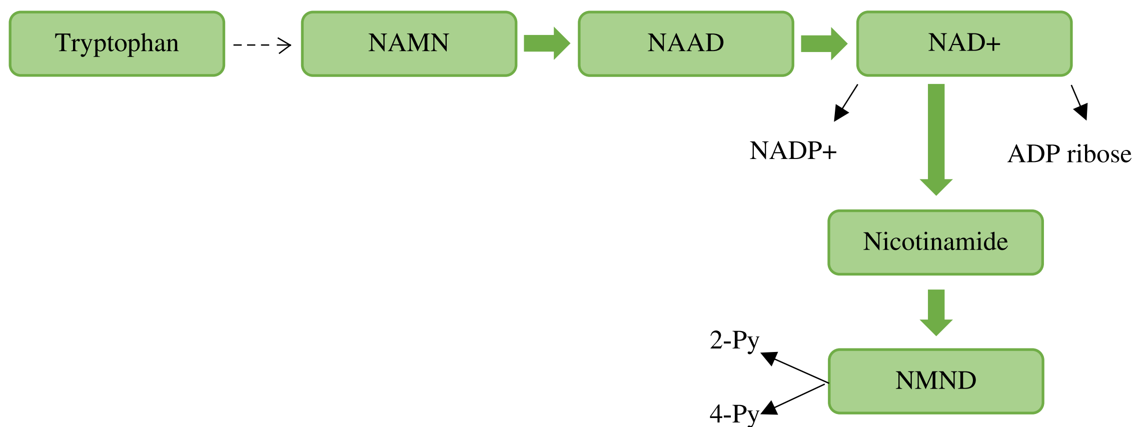


Figure 16 – Schematic representation of metabolic pathway of Nicotinate and Nicotinamide. NAMN- nicotinic acid mononucleotide; NAAD- nicotinic acid adenine dinucleotide; NAD- nicotinamide adenine dinucleotide; NADP-nicotinamide adenine dinucleotide phosphate.

Lung surfactant production

Variations in ethanolamine (decreased) and *m*-inositol (increased) were found in PTB NBs. These metabolites are precursors of phosphatdylethanolamine and phosphadidylinositol respectively and act as lung surfactants. In this study ethanolamine was found decreased suggesting an alteration on surfactants formations due to gestational age that these NB born (31-32). It was reported that phosphatidylinositol productions increase from the 28th week with maximum values between week 32 and 35 and declining until the end of gestation.

Previous reports had related that altered renal function may increase *m*-inositol in urine of PTB NB (57). According to Dessi et al 2012 *m*-inositol was found increased in IUGR NB possibly due insulin resistance. As said previously, these 2 PTB NBs born at 31 and 32 g.a., thus its increase may be a reflection of higher phosphatidylinositol formation and/or a consequence of altered renal function.

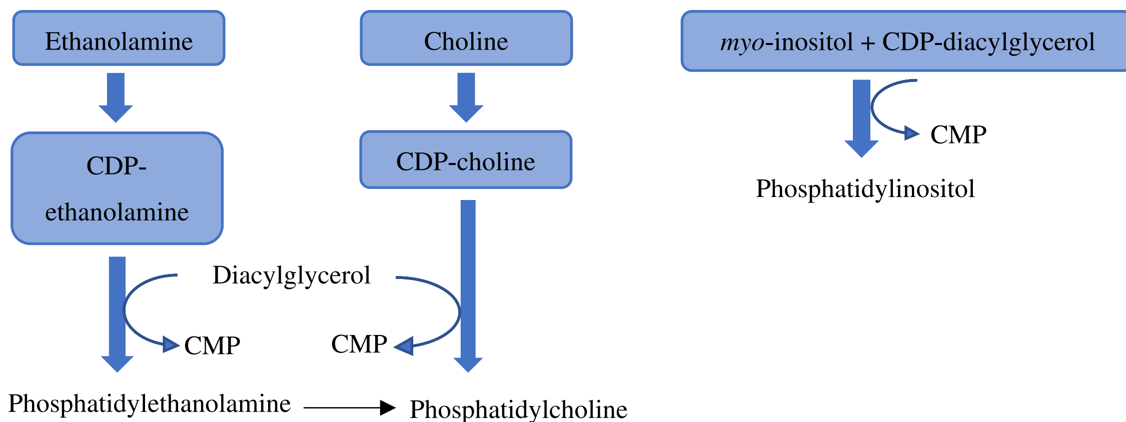


Figure 17 – Biosynthesis of phosphatidylethanolamine, phophatidylcholine and phosphatidylinositol. CDP- cytidine diphosphate; CMP- cytidine monophosphate.

Renal function

Creatine is important in muscle metabolism and creatinine is formed by breakdown product of creatine phosphate in muscle. The metabolite was found decreased in PTB NBs. According to creatinine urinary and blood concentration may be altered by age, weigh, gender and is a marker of glomerular of filtration. Lower creatinine in PTB NB urine has been reported to immature tubular and vascular structures in the kidney.

At the time of the birth, betaine concentration on PTB NBs is higher possibly due to immature kidneys. Choline is a precursor of phophatidylcholine and the syntheses of this precursor increased from week 28th until the end of gestation with peak production at 36th g.w. Choline is also a precursor of betaine and it is had been identified as age-dependent in NBs and adults. In this study betaine was found decreased.

DMA had been related with gut microflora and was found decreased. The reason for this variation is that the gut microflora of PTB NBs is poor. DMG was found increased and it is possible related with alterations on choline metabolism. Alterations on methionine metabolism is consistent with increased DMG.

The increase of taurine had been previously reported (53,57). A study has reported that variation in taurine in NBs with BDP may be involved in renal damage, osmoregulation and fetal wellbeing.

Other metabolic effects

Galactose was found increased in PTB NBs, possibly related with breast or formula milk feeding, since lactose is the major sugar present in these NBs feedings and source of glucose on galactose.

Hippurate, 1-methylhistidine and histidine are related with gut microflora and in this study were found decreased and increased respectively. Previously reports had shown that hippurate increased with age and with consumption of phenolic compounds (these molecules are produced by gut microbiota).

4.3. Statistical analysis of newborn urine spectra at birth (without outliers)

The previous subchapter shows visual alterations of spectra per class and the average changes related to bubble size PLS-DA and because there are different stages of prematurity it is necessary to proceed with pairwise multivariate analysis to identify the changes in urine, in each level of development.

4.3.1. Extremely Preterm + Very Preterm vs Controls

In order to evaluate the impact of different stages of PTB in the urinary metabolome in NBs, PCA and PLS-DA were applied. Firstly, EPT was compared with controls and scores plot of PCA revealed an obvious separation trend between groups (along PC1) (*Figure 18* a)). Due to the reduced number of samples of EPT (n=2), no further statistical analyses could be done for this group. Then EPT was compared with VPT and scores plot of PCA did not reveal a separation trend (*Figure 18* b)). This group has a reduced number of samples to VPT (n=5), so that again no further statistical analysis was done. Subsequently both EPT and VPT groups were joined and studied as one single group. The analysis of PCA scores plot of EPT/VPT compared with controls showed a separation between groups, later confirmed by PLS-DA (*Figure* 18 c), d)). The PLS-DA presents an apparently high predictive power (Q^2 of 0,877) and, in spite of the low sample numbers of premature newborns. The analysis of the loading plot obtained from PLS-DA unveiled differences in some metabolites (indicated in *Figure 19*). The contribution of metabolites for the separation of groups is represented by colour, which means that red peaks are those with high contribution and blue peaks are those with low contribution. Also, positive peaks on loadings are those that are in positive part of PLS-DA, and negative peaks are those in negative part of PLS-DA.

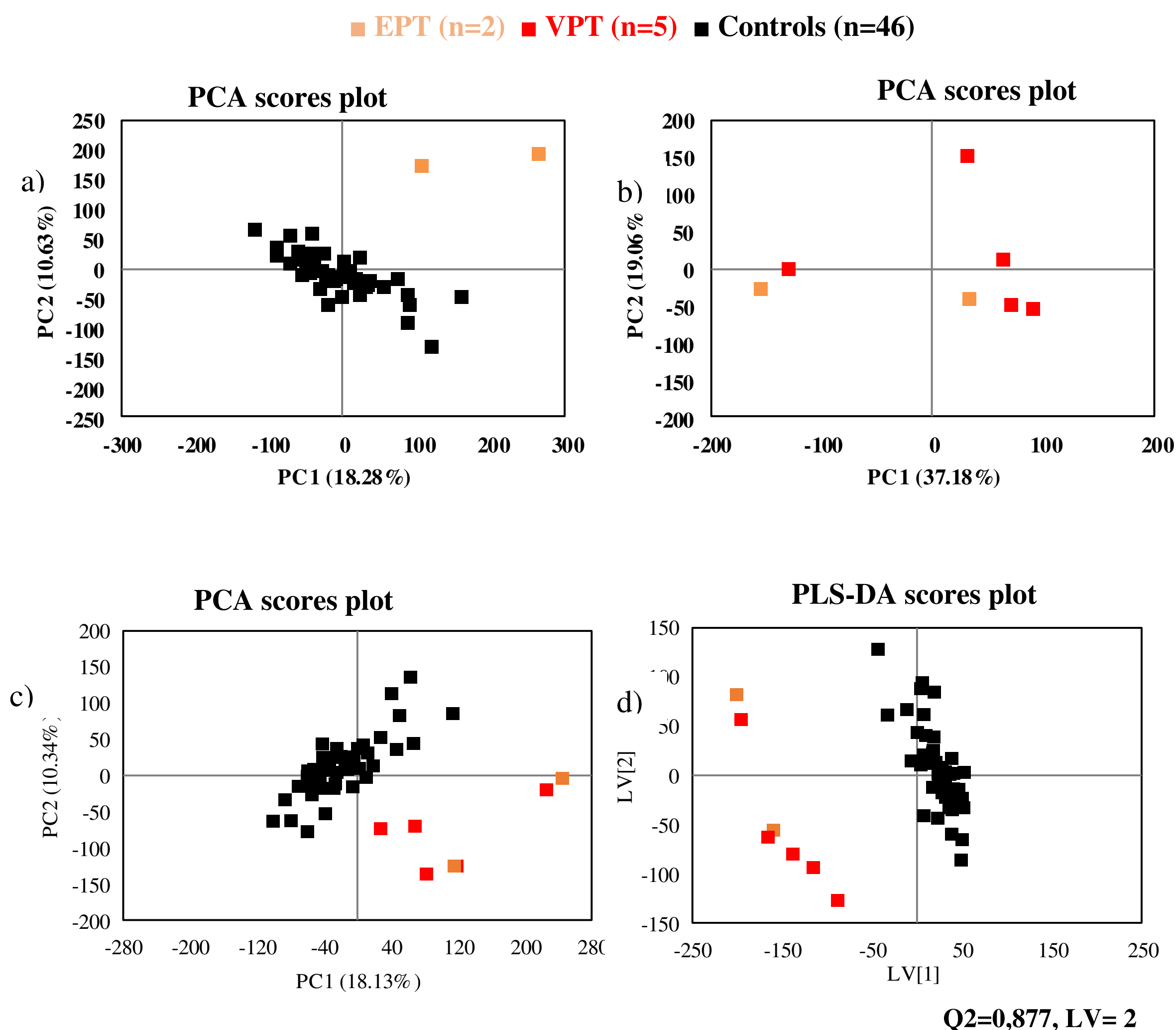


Figure 18 – a) PCA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ EPT) and controls (■); b) PCA scores scatter plot of ^{1}H NMR spectra of urine between PTB newborns (■ **EPT and** ■ **VPT**); c) PCA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ EPT **plus** ■ **VPT**) and controls (■); d) PLS-DA scores scatter plot of ^{1}H NMR spectra of urine from PTB newborns (■ EPT **plus** ■ **VPT**) and controls (■).

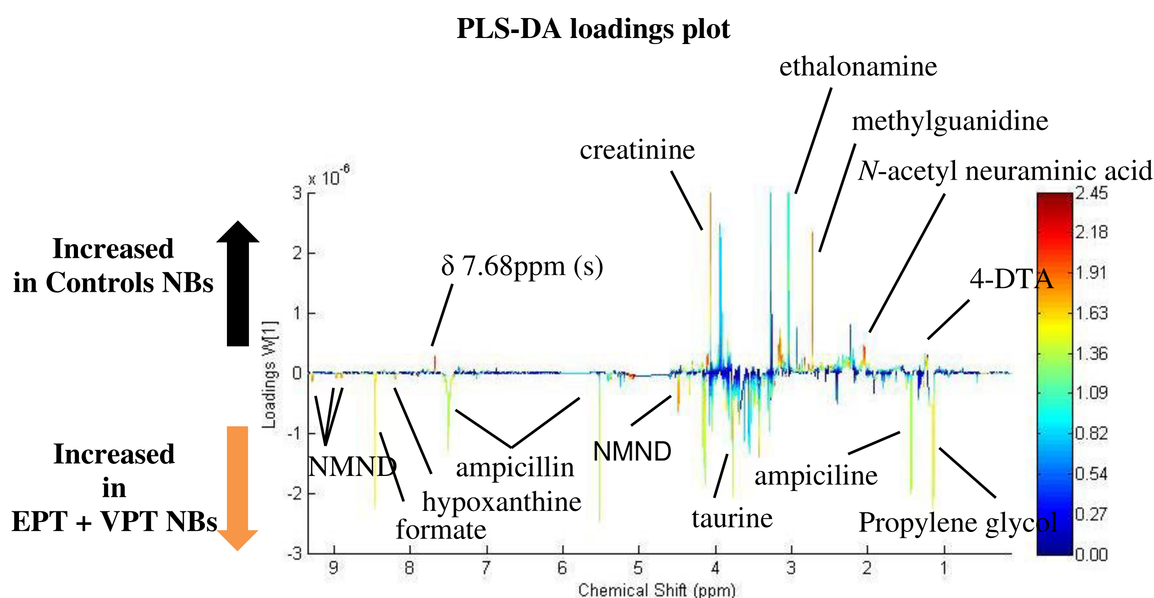


Figure 19 – PLS-DA loadings plots of ^{1}H NMR spectra of urine from PTB newborns (**EPT plus VPT**) and controls, with some peak assignments.

Subsequently, the peaks contributing more strongly were further evaluated by spectral integration and univariate statistical analysis. The metabolic changes with higher contributions were analysed in terms of magnitude of variation (effect-size) and statistical significance (*p-value*), furthermore tested with Bonferroni and the FDR test. Only NMND variations were found to remain statistically significant upon the strict Bonferroni test ($1{,}47 \times 10^{-03}$), whereas more variation remained relevant upon the FDR test. Hence, *Table 6* lists the results obtained for effect-size and *p*-value obtained with the latter test. In this way, relevant changes between EPT/VPT and Controls were found for 9 metabolites (NMND, formate, hypoxanthine, 1-methylhistidine, creatinine, ethanolamine, *cis*-aconitate, methylguanidine and 3-HBA), 4 unassigned compounds (U2 (7,88 ppm), U3 (7,68 ppm), U16 (3,09 ppm) , U18 (2,78 ppm), and 2 unassigned spectral regions (Region1 (2.78-2.77 ppm), Region2 (1,25-1,23 ppm)).

Table 6 - Metabolite/resonance changes in PTB NBs urine. Legend: s: singlet, d: doublet, t: triplet, m: multiplet, Ui: unassigned compound i by order of appearance, NMND: *N*-methyl nicotinamide, 2-Py: *N*-mehyl-2-pyridone-5-carboxamide, 3-HBA: 3-hydroxybutyrate. Only p-values<0.05 are shown and was apply FDR test.

EPT (n=2) + VPT (n=5) *vs* Controls (n=46)		
Metabolite	**δ_H/ppm and multiplicity**	**Variation (effect size; p-value; p-value with FDR)**
NMND	4,48 (s); 8,90 (d); 8,97 (d); 9,27 (s)	↑ (2,45; $1{,}47 \times 10^{-03}$; $9{,}19 \times 10^{-05}$)
formate	8,46 (s)	↑ (2,04; $4{,}60 \times 10^{-02}$; $4{,}31 \times 10^{-02}$)
2-Py	6,67 (d); 7,97 (dd); 8,33 (d)	↓ (-2,29)
hypoxanthine	8,20	↓ (-0,88; $4{,}60 \times 10^{-02}$; $4{,}60 \times 10^{-02}$)
4-OH-hyppurate	3,96 (s);6,98 (d); 7,76 (d)	↓ (-1,81)
1-methylhistidine	3,16 (dd); 3,72 (s); 3,97 (dd); 7,04 (s); 7,78 (s)	↓ (-2,79; $2{,}27 \times 10^{-02}$; $1{,}78 \times 10^{-02}$)
sucrose	5,41 (d)	↑ (0,97)
glucose	4,65 (d); 3,43 (t); 3,54 (dd); 5,24 (d)	(0,80; $2{,}59 \times 10^{-02}$; $-$; $1{,}78 \times 10^{-02}$)
galactose	3,64 (dd); 3,75 (m); 3,83 (m); 3,99 (d); 4,07 (t); 4,59 (d); 5,28 (d)	↓ (-2,28)
creatinine	3,05 (s); 4,06 (s)	↓ (-2,79; $1{,}60 \times 10^{-02}$; $6{,}02 \times 10^{-03}$)
m-inositol	3,29 (t); 3,55 (dd); 3,63 (t); 4,07 (t)	↑ (1,36)
taurine	3,26 (t); 3,43 (t)	↑ (1,64)
ethanolamine	3,15 (t); 3,83 (t)	↓ (-2,50; $2{,}12 \times 10^{-02}$; $1{,}06 \times 10^{-02}$)
cis-aconitate	3,12 (d)	↓ (-2,84; $1{,}50 \times 10^{-02}$; $4{,}67 \times 10^{-03}$)
methylguanidine	2,83 (s)	↓ (-2,59; $2{,}27 \times 10^{-02}$; $1{,}27 \times 10^{-02}$)
DMA	2,72 (s)	↓ (-3,09)
N-acetylneuraminic acid	2,05 (d)	↓ (-2,76)
alanine	1.49 (d); 3,78 (q)	↑ (0,98)
3-HIVA	1,27 (d)	↑ (2,39)
4-DTA	1,23 (d)	↓ (-2,08)
propylene glycol	1,15 (d)	↑ (1,94)
U1	7,90	↑ (3,49)

U2	7,88	↑ (1,36; $1,85 \times 10^{-2}$; $-$; $8,08 \times 10^{-3}$)
U3	7,68	↓ (-3,78; $4,60 \times 10^{-02}$; $4,02 \times 10^{-02}$)
U4	5,37	↑ (2,41)
U5	5,10	↑ (3,40)
U6	5,08	↑ (4,46)
U7	4,33	↑ (1,57)
U8	4,29	↓ (-1,62)
U9	4,16	↑ (1,49)
U10	4,14	↑ (1,63)
U12	3,77	↑ (1,72)
U13	3,39	↑ (2,49)
U14	3,23	↑ (1,16)
U15	3,22	↑ (3,15)
U16	3,09	↓(-0,98; $3,82 \times 10^{-02}$; -; $2,86 \times 10^{-2}$)
U17	2,90	↓ (-3,07)
U18	2,78	↓(-2,93;8.95×10^{-03}; $1,12 \times 10^{-3}$)
U19	2,76	↓ (-3,66)
U20	1,43	↑ (1,65)
Region1	[2,78-2,77]	↓ (-3,23; $1,04 \times 10^{-02}$; $-$; $1,95 \times 10^{-03}$
Region2	[1,25-1,23]	↓ (-2,17; $4,06 \times 10^{-02}$; $-$; $3,30 \times 10^{-02}$)

Proposed metabolic interpretation of changes related with prematurity (EPT/VPT vs Controls)

In this cohort metabolic changes were observed in NMND (involved in nucleotide metabolism), formate (involved in purine biosynthesis), hypoxanthine (involved in purine metabolism), 1-methylhistidine, creatinine (possible related with renal function or oxidative stress), ethanolamine (precursor of lung surfactant phosphatidylethanolamine), *cis*-aconitate (involved in TCA cycle), methylguanidine and 3-HBA. This means that the signature registered may be indicative of underdevelopment in certain organs such lungs (since these are the last organs to develop), heart, brain and intestines.

4.3.2. Extremely Preterm/Very Preterm vs Moderate to Late Preterm

For this study PCA was applied to compare EPT/VPT with MLPT and it is not possible to observe a generally separation between groups (*Figure 20* a)) which means that changes are very small in number and/ magnitude. When PLS-DA is applied (*Figure 20* b)),

some group separation is more notable. However, the PTB13 individual is more deviated from the remaining MLPT group, which means that its classification may be intermediate between the two groups compared here. In terms of gestational age, PTB13 is distant from EPT/VPT since his born at 35 g.w., then is possible that this PTB is a deviant profile which may be interesting to investigate in future. So for PLS-DA scores scatter plot this spectrum was removed in order to obtain a better model with good separation between groups (*Figure 20* c)). The analysis of the loading plot unveiled differences in some metabolites (indicated in *Figure 21*), these being further evaluated by spectral integration and univariate comparison.

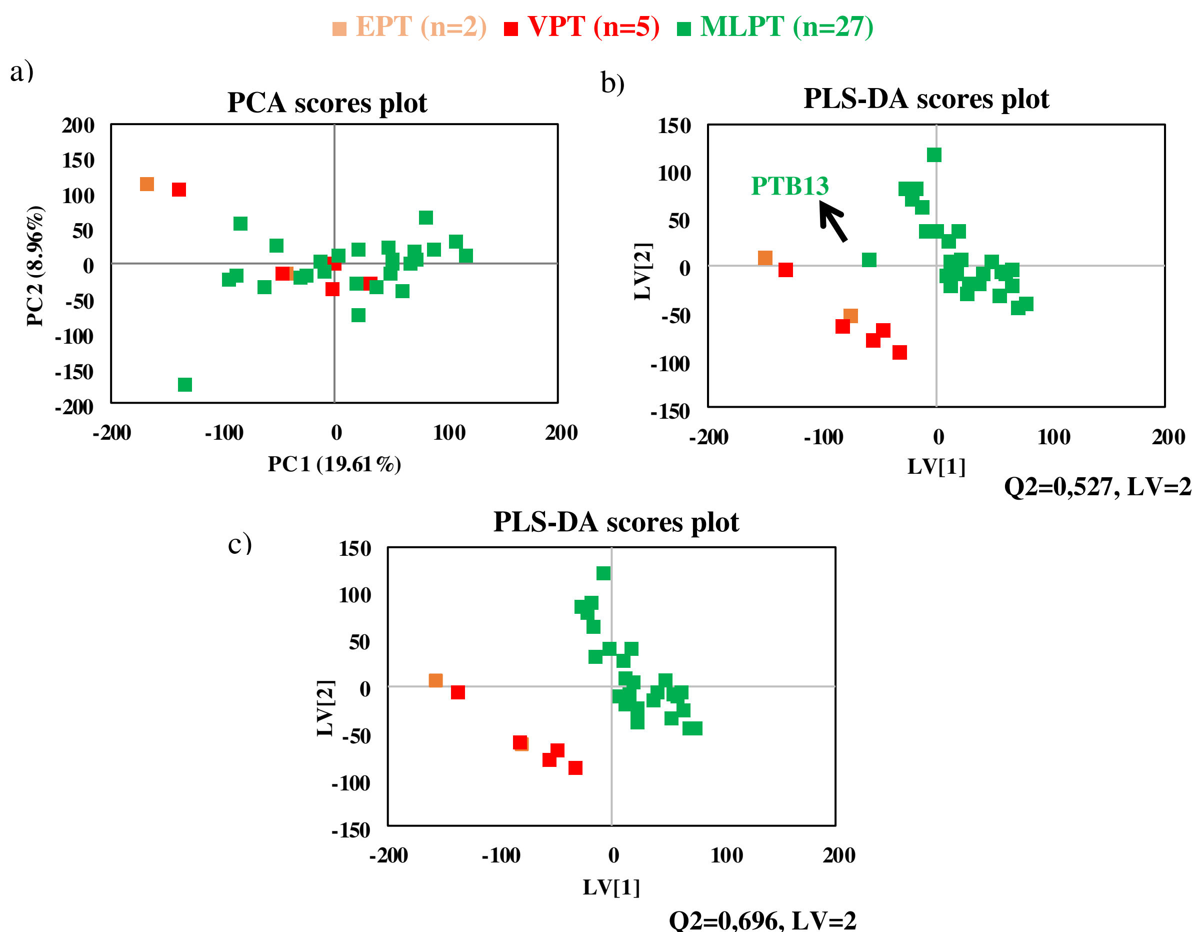


Figure 20 – a) PCA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ EPT **plus** ■ VPT) and MLPT (■); b) PLS-DA scores scatter plot of ^{1}H NMR spectra of urine from PTB newborns (■ EPT **plus** ■ VPT) and MLPT (■); c) PLS-DA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ EPT **plus** ■ VPT) and MLPT (■) without RN13.

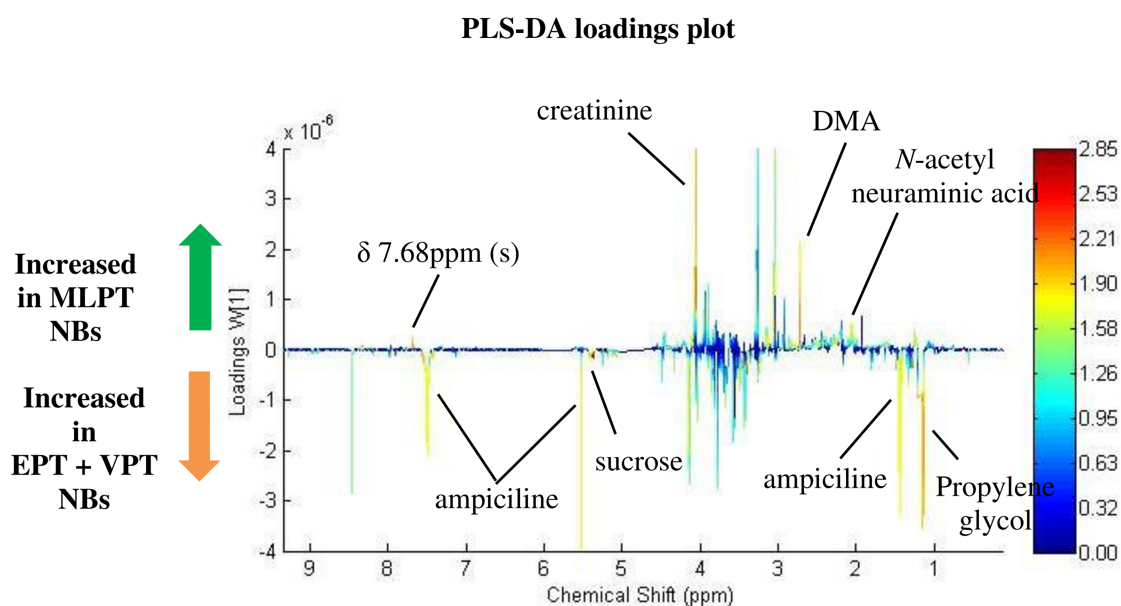


Figure 21 – PLS-DA loadings plots of ^{1}H NMR spectra of urine from PTB newborns (**EPT plus VPT**) and **MLPT**, with some peak assignments.

These metabolic changes were analysed in terms of biological (effect-size) and statistical significance (*p-value*). Was not found statistical significance for any metabolite since p-value>0.05 (using both Bonferroni and FDR test). The variations found are listed in *Table 7*.

Table 7 **-** Metabolite/resonance changes in PTB NBs urine. Legend: s: singlet, d: doublet, t: triplet, m: multiplet, Ui: unassigned compound i by order of appearance, NMND: N-methyl nicotinamide, 2-Py: N-mehyl-2-pyridone-5-carboxamide, 3-HBA: 3-hydroxybutyrate. Only p-values<0.05 are shown and was apply FDR test.

EPT (n=2) + VPT (n=5) *vs* MLPT (n=26)		
Metabolite	**δ H/ppm and multiplicity**	**Variation (effect size)**
NMND	4,48 (s); 8,90 (d); 8,97 (d); 9,27 (s)	↑ (0,85)
formate	8,46 (s)	↑ (0,98)
2-Py	6,67 (d); 7,97 (dd); 8,33 (d)	↓ (-1,66)
4-OH-hyppurate	3,96 (s); 6,98 (d); 7,76 (d)	↓ (-1,38)
sucrose	5,41 (d)	↑ (1,14)
galactose	3,64 (dd); 3,75 (m); 3,83 (m); 3,99 (d); 4,07 (t); 4,59 (d); 5,28 (d)	↓ (-1,73)
creatinine	3,05 (s); 4,06 (s)	↓ (-1,64)
ethanolamine	3,15 (t); 3,83 (t)	↓ (-1,09)
cis-aconitate	3,12 (d)	↓ (-1,37)
methylguanidine	2,83 (s)	↓ (-1,21)
DMA	2,72 (s)	↓ (-1,43)
N-acetylneuraminic acid	2,05 (d)	↓ (-1,07)
3-HIVA	1,27 (d)	↑ (1,63)
4-DTA	1,23 (d)	↓ (-1,27)
Propylene glycol	1,15 (d)	↑ (1,48)
U1	7,90	↑ (2,31)
U2	7,88	↑ (1,10)
U3	7,68	↓ (-1,49)
U4	5,37	↑ (2,39)
U5	5,10	↑ (1,02)
U6	5,08	↑ (1,16)
U8	4,29	↓ (-0,93)
U9	4,16	↑ (1,28)
U10	4,14	↑ (0,86)
U13	3,39	↑ (0,92)
U14	3,23	↑ (1,37)
U15	3,22	↑ (1,10)
U17	2,90	↑ (1,23)
U18	2,78	↑ (1,23)
U19	2,76	↑ (1,58)
U20	1,43	↑ (1,26)

In this group study do not was obtain significant metabolic changes and these can be explained by the small cohorts of two groups.

4.3.3. Moderate to late Preterm vs Controls

Lastly, MLPT was compared with controls and scores scatter plot of PCA shown a separation trend between groups with some overlap confirmed by PLS-DA scores scatter plot (*Figure 22* a),b)). The analysis of the loading plot unveiled differences in some

metabolites (indicated in *Figure 23*), these being further evaluated by spectral integration and univariate comparison.

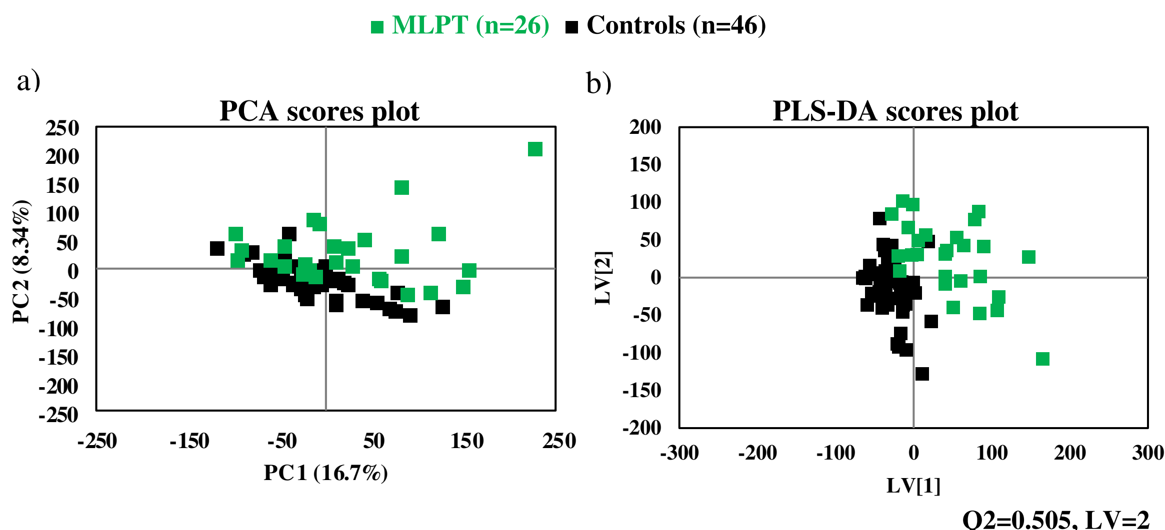


Figure 22 – a) PCA scores scatter plots of ^{1}H NMR spectra of urine from PTB newborns (■ MLPT) and **Controls** (■); b) PLS-DA scores scatter plot of ^{1}H NMR spectra of urine from PTB newborns (■ MLPT) and **Controls** (■).

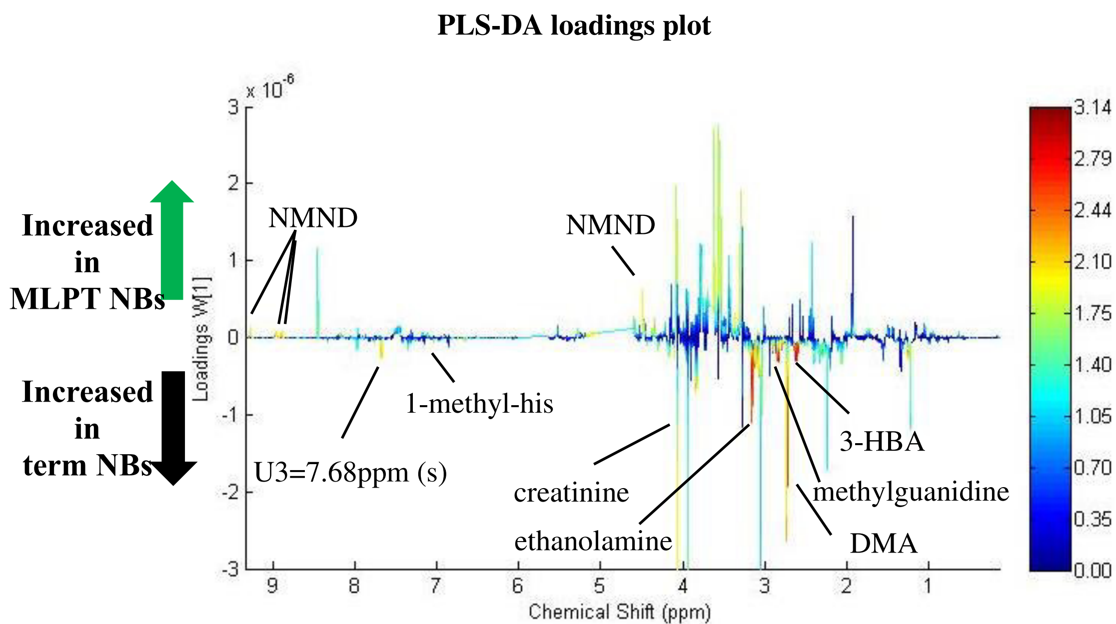


Figure 23 – PLS-DA loadings plots of ^{1}H NMR spectra of urine from PTB newborns (MLPT) and **Controls**, with some peak assignments.

These metabolic changes were analysed in terms of effect-size and statistical significance (p-value) and listed in *Table 8* with the corresponding variations. Relevant changes were found for 10 metabolites (NMND, 2-Py, 1-methylhistidine, creatinine, ethanolamine, *cis*-aconitate, methylguanidine, DMA, 3-HBA, *N*-acethyl neuraminic acid, 4-DTA), 10 unassigneds compounds (U3 (7,68 ppm), U5 (5,10 ppm), U6 (5,08 ppm), U7 (4,33 ppm), U12 (3,77 ppm), U13 (3,39 ppm), U15 (13,22 ppm), U16 (3,09 ppm), U17 (2,90 ppm), U18 (2,78 ppm), U19 (2,76 ppm)) and 2 unassigned spectral regions (Region1 (2,78-2,77 ppm), Region2 (1,25-1,23 ppm)).

Table 8 - Metabolite/resonance changes in PTB NBs urine. Legend: s: singlet, d: doublet, t: triplet, m: multiplet, Ui: unassigned compound i by order of appearance, NMND: N-methyl nicotinamide, 2-Py: N-mehyl-2-pyridone-5-carboxamide, 3-HBA: 3-hydroxybutyrate. Only p-values<0.05 are shown and was apply FDR test.

MLPT (n=26) vs Controls (n=46)		
Metabolite	**δ_H/ppm and multiplicity**	**Variation (effect size; p-value; Bonferroni; FDR)**
NMND	4,48 (s); 8,90 (d); 8,97 (d); 9,27 (s)	↑ (1,02; $6{,}41 \times 10^{-05}$; $6{,}41 \times 10^{-05}$; $1{,}33 \times 10^{-05}$)
2-Py	6,67 (d); 7,97 (dd); 8,33 (d)	↓ (-0,82; $1{,}48 \times 10^{-02}$; $-$; $1{,}29 \times 10^{-02}$)
1-methylhistidine	3,16 (dd); 3,72 (s); 3,97 (dd); 7,04 (s); 7,78 (s)	↓ (-1,05; $5{,}42 \times 10^{-03}$; $-$; $3{,}39 \times 10^{-03}$)
creatinine	3,05 (s); 4,06 (s)	↓ (-1,10; $1{,}28 \times 10^{-03}$; $1{,}28 \times 10^{-03}$; $6{,}38 \times 10^{-04}$)
ethanolamine	3,15 (t); 3,83 (t)	↓ (-1,42; $1{,}16 \times 10^{-03}$; $1{,}16 \times 10^{-03}$; $5{,}32 \times 10^{-04}$)
cis-aconitate	3,12 (d)	↓ (-1,65; $2{,}11 \times 10^{-06}$; $2{,}11 \times 10^{-06}$; $8{,}79 \times 10^{-08}$)
methylguanidine	2,83 (s)	↓ (-1,13; $6{,}61 \times 10^{-03}$; $-$; $4{,}68 \times 10^{-03}$)
DMA	2,72 (s)	↓ (-1,25; $1{,}88 \times 10^{-04}$; $1{,}88 \times 10^{-04}$; $5{,}48 \times 10^{-05}$)
N-acetylneuraminic acid	2,05 (d)	(-0,57; $6{,}22 \times 10^{-03}$; $-$; $4{,}15 \times 10^{-03}$)
3-HBA	1,20 (d); 2,31 (m); 2,41 (m); 4,15 (m)	↓ (-1,72; $1{,}85 \times 10^{-06}$; $1{,}85 \times 10^{-06}$; $1{,}54 \times 10^{-07}$)
4-DTA	1,23 (d)	↓ (-0,88; $3{,}33 \times 10^{-03}$; $-$; $1{,}94 \times 10^{-03}$)
U3	7,68	↓ (-1,07; $1{,}43 \times 10^{-03}$; $1{,}43 \times 10^{-03}$; $7{,}74 \times 10^{-04}$)
U7	4,33	↑ (0,88; $1{,}28 \times 10^{-02}$; $-$; $1{,}02 \times 10^{-02}$)
U16	3,09	↓ (-1,16; $1{,}99 \times 10^{-04}$; $1{,}99 \times 10^{-04}$; $4{,}96 \times 10^{-05}$)
U17	2,90	↓ (-1,51; $7{,}17 \times 10^{-05}$; $7{,}17 \times 10^{-05}$; $1{,}19 \times 10^{-05}$)
U18	2,78	↓ (-1,15; $2{,}71 \times 10^{-04}$; $2{,}71 \times 10^{-04}$; $9{,}04 \times 10^{-05}$)
U19	2,76	↓ (-1,30; $2{,}86 \times 10^{-04}$; $2{,}86 \times 10^{-04}$; $1{,}07 \times 10^{-04}$)
Region1	[2,78-2,77]	↓ (-1,16; $9{,}02 \times 10^{-05}$; $-$; $1{,}13 \times 10^{-05}$)
Region2	[1,25-1,23]	↓ (-0,71; $6{,}93 \times 10^{-03}$; $-$; $5{,}20 \times 10^{-03}$)

Analyzing both data of metabolites obtained from loadings and from *Table 8*, it is possible observe that those with more relevance on loadings remains significant. Comparing the

results obtain in this study (EPT/VPT vs MLPT) with previous study (EPT/VPT vs Controls), a greater number of significant metabolites was found, such as, 2-Py, DMA, *N*-acethyl neuraminic acid, 4-DTA, U5, U6, U7, U8, U12, U13, U15, U17 and U19. Althougt, formate, hypoxanthine and U2 only remains significant on EPT/VPT vs Controls, indicating possible relation of these metabolites with lower stages of prematurity.

Proposed metabolic interpretation of changes related with prematurity (MLPT vs Controls)

In this cohort we have observed metabolic changes in NMND (involved in nucleotid metabolism), 2-Py, 1-methylhistidine (related with gut microflora), creatinine (possible related with renal function or oxidative stress, ethanolamine (precursor of lung surfactant phosphatidylethanolamine), *cis*-aconitate (involved in TCA cycle), methylguanidine, DMA, 3-HBA, *N*-acetyl neuraminic acid and 4-DTA.

In this study a greater number of significant metabolites was found (comparatively to previous study), such as 2-Py, DMA, *N*-acetyl neuraminic acid, 4-DTA, U5, U6, U7, U8, U12, U13, U15, U17 and U19. This means that the signature registered may be indicative of late stages of prematurity.

4.3.4. Effects of different stages of prematurity on urine composition: proposed metabolic interpretation

Analyzing the results obtained previously on visual comparison between different stages of prematurity and controls, is observed that alterations on NMND, formate, DMA and ethanolamine, found in all stages, were found significant. Creatine alterations found in EPT and VPT group were found significant in all studies. Furthermore, visual inspection of the spectra of EPT shows alterations on 2-Py, histidine and fumarate, however in PCA and PLS-DA results between groups, formate, only remain significant on EPT/VPT vs Controls. No modifications were detected on hypoxanthine and U2 on visual spectra, however these two metabolites are significative on EPT/VPT vs Controls, indicating possible relation of these metabolites with lower stages of prematurity. Metabolites such as 2-Py, DMA, *N*-acetyl neuraminic acid, 4-DTA, U5 (5,10 ppm), U6 (5,08 ppm), U7 (4.33ppm), U12 (3,77 ppm), U13 (3,39 ppm), U15 (3,22 ppm), U16 (3.09ppm), U17 (2.90ppm), U18 (2.78ppm),

U19 (2.76ppm) were found significant only on MLPT vs Controls, meaning that the signature registered may be indicative of late stages of prematurity. 1-methylhistidine, *cis*-aconitate, methylguanidine and 3-HBA were found significant too.

Concluding, formate, hypoxanthine and U2 (7,88 ppm) may be related with lower stages of prematurity, while 2-Py, DMA, N-acetyl neuraminic acid, 4-DTA, U5 (5,10 ppm), U6 (5,08 ppm), U7 (4.33ppm), U12 (3,77 ppm), U13 (3,39 ppm), U15 (3,22 ppm), U17 (2.90ppm) and U19 (2.76ppm) may be related with later stages of prematurity. Since NMND, 1-methylhistidine, creatinine, ethanolamine, cis-aconitate, methylguanidine, 3-HBA, U3 (7.68ppm), U16 (3.09ppm), U18 (2.78ppm), Region1 (2.7885-2.7738) and Region2 (1.2505-1.2366ppm) where found significant for all groups, may indicate that are relative with PTB.

Chapter 5. Longitudinal study of individual newborns with different prematurity degrees – Study 2

This chapter describes the results obtained for the NMR study of PTB NB urinary trajectory in order to characterize the metabolic changes occurring during the NBs hospital stay (*Study 2*). For this study more than one sample were collected for each NB in different days. The following section shows the PCA scores scatter plots of all of these PTB newborns, representing each individual by a different color so that the corresponding trajectory may be easily identified (*Figure 24*).

5.1. Statistical analysis of newborns urine spectra and metabolites trajectories

The analysis of PCA scores plot (*Figure 24*) shows that, in general and as expected, PTB NBs tend to approach the urine profile corresponding to term NBs over time. However, it is possible to note some exceptions to this tendency. For instance, PTB1 practically does not vary, while PTB2, 4, 6 and 7 approach to controls. In PTB3 is possible to see a metabolic alteration after 4 days in the hospital (making it move away from the controls), which is then recovered the next day. PTB5 is one where more instability is visible during your stay in hospital. At day 3 is visible a metabolic alteration (making it move away from the controls), which is then recovered in days 4 and 5. At days 9 and 17 is visible again a metabolic alteration, recovered in day 19. New metabolic alteration at day 23 and recovered in the next day. In PTB9 is visible a metabolic alteration, however, is difficult to conclude about metabolic changes since there are only 2 collection samples (1st collection corresponding to sample RN17 (that was excluded from VPT group) analysed in Chapter 4. Lastly, in PTB10 is visible an approach to controls between day 1 to 6, whilst in day 14 is visible a metabolic alteration making it move away from the controls. Thus, since PTB3, PTB5, PTB9 and PTB10 seem to behave deviant from controls, a personalized monitoring is important.

■ Controls (n=46) ● PTB1 (n=2) ● PTB2 (n=3) ● PTB3 (n=3) ● PTB4 (n=2) ● PTB5 (n=9)
● PTB6 (n=2) ● PTB7 (n=3) ● PTB8 (n=1) ● PTB9 (n=2) ● PTB10 (n=3)

PCA scores scatter plot

PC2 (8%)

PC1 (19,66%)

Figure 24 – PCA scores scatter plots of ^{1}H NMR spectra of ten PTB NBs (all the coloured circles) trajectory urinary and **Controls** (■).

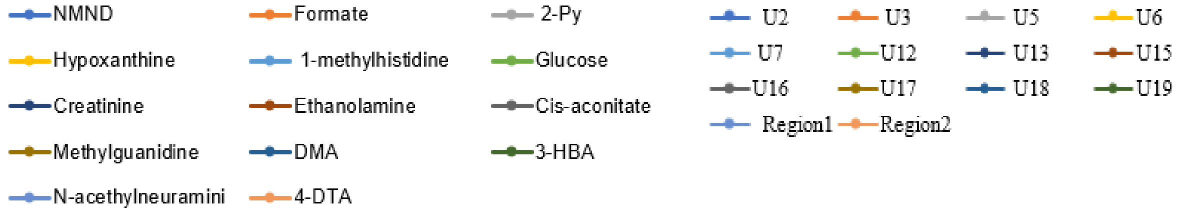
In *Table 9* are the list of PTB NBs grouped by colours (dark grey is the PTB NBs closely from controls and light grey are the furthest), the variation seen in the PCA scores plot, respective g.a. clinical information and g.a. at collection.

Table 9 – List of PTB NBs grouped by colours (dark grey is the PTB NBs closely from controls and light grey are the furthest), the variation seen in the PCA scores plot, respective g.a. and clinical information.

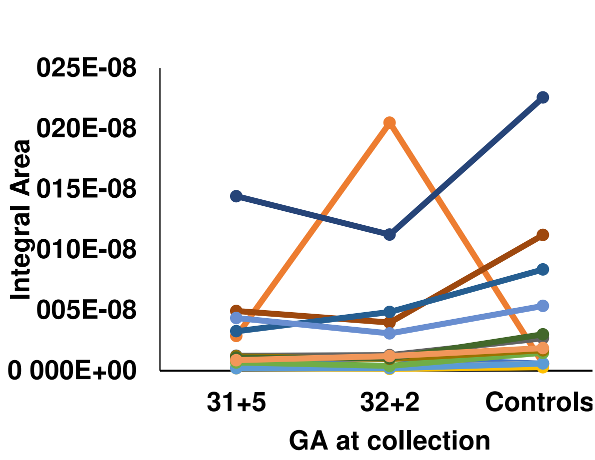
PTB	Controls / Variation	g.a.	Clinical Information
1	Close / Stable	31 (VPT)	*; VD; M
2	Close / Stable (approach to controls)	35 (MLPT)	Born due of membrane rupture; tachypnea transitory of NB, conjutivity, hyperbilurubinemia; VD; M
3	Distant / Unstable	33 (MLPT)	Mother with previously abortion; *; CS; M
4	Close / Stable (approach to controls)	31 (VPT)	Mother with PRE-eclampsia; *: CS; F
5	Distant / Unstable	26 (EPT)	Risk of infection, sepsis, hyponatremia, hyperbilirubinemia and Mother with 6 previous gestations has gestational diabetes VD; F
6	Distant / Stable (approach to controls)	35 (MLPT)	Tachypnea transitory of NB; pneumothorax; hyperbilurubinemia; twin; VD; F
7	Close / Stable (approach to controls)	34 (MLPT)	Hyperbilurubinemia; twin of PTB8; CS; F
8	Close / Without variation because it is only 1 sample	34 (MLPT)	*; twin of PTB7; CS; M
9	Distant / Move away from controls	31 (VPT)	Tachypnea transitory of NB; twin of PTB10; VD; M
10	Distant / Unstable	31 (VPT)	Hyalines membranes disease; twin of PTB9; VD; M

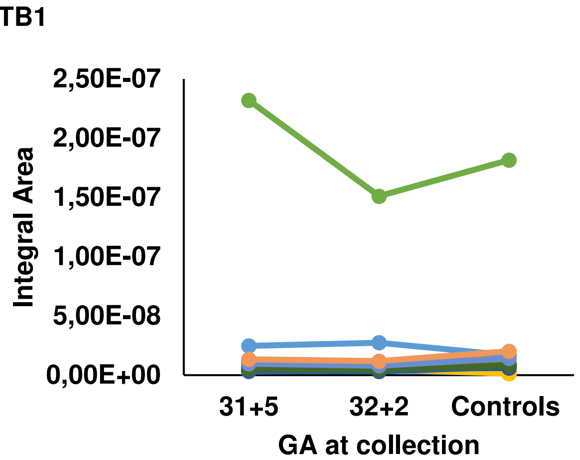
*Without complications after birth

In order to see the metabolic variations for each PTB over time, was made a trajectory of metabolites found with significant variance in Chapter 4 and compared with found in controls (*Figure 25 and 26*).

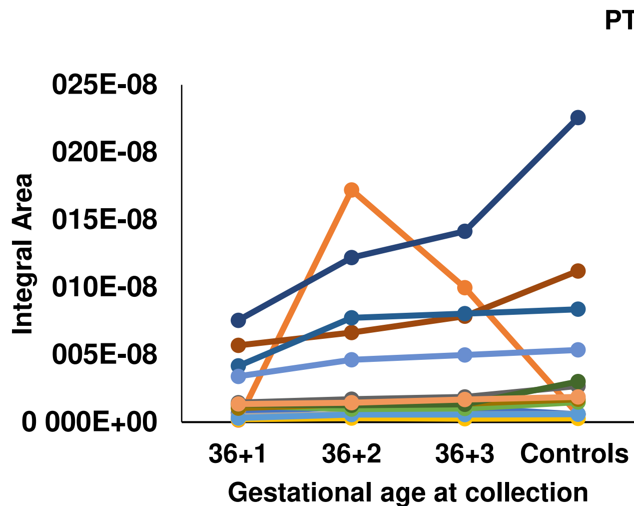


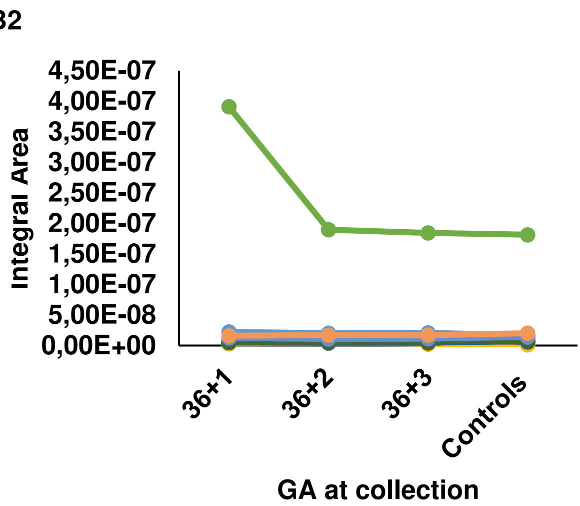
PTB1



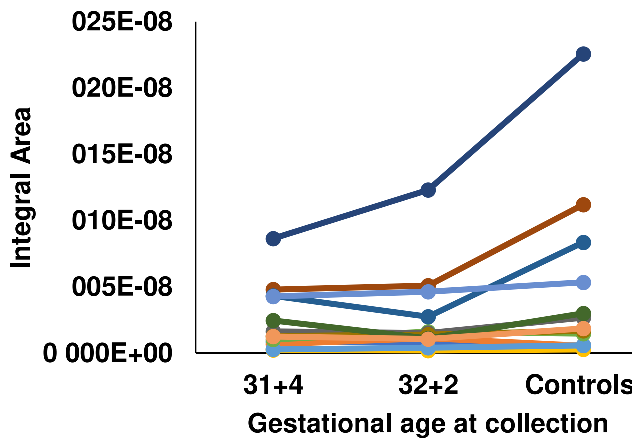


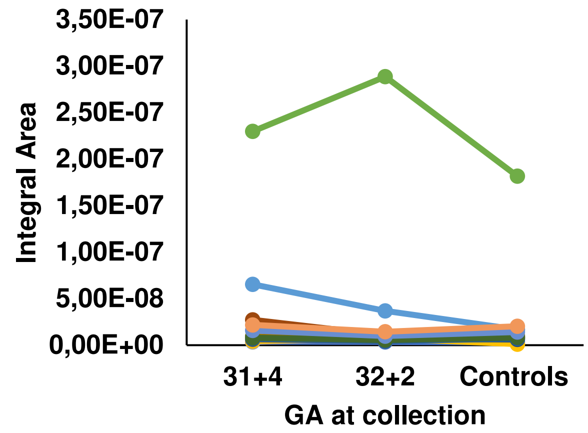
PTB2



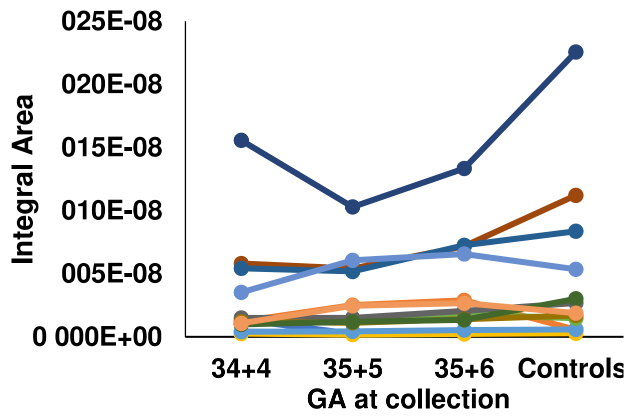


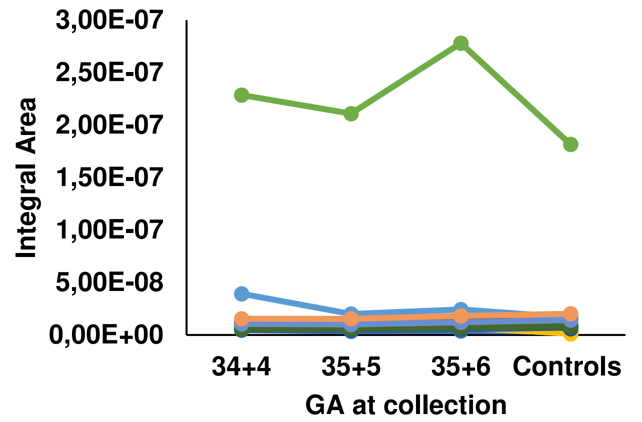
PTB4





PTB7





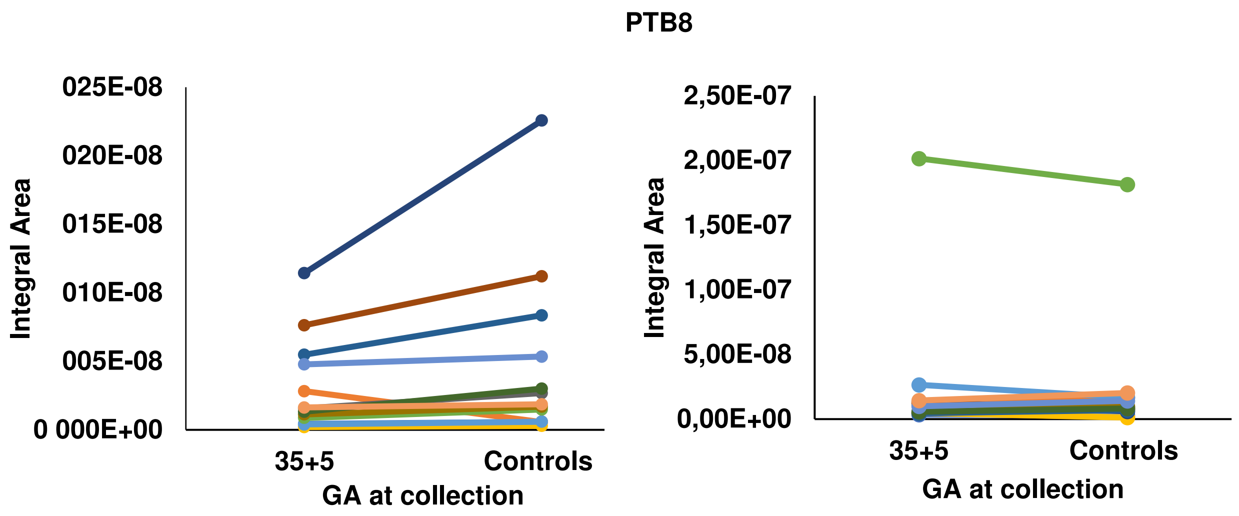
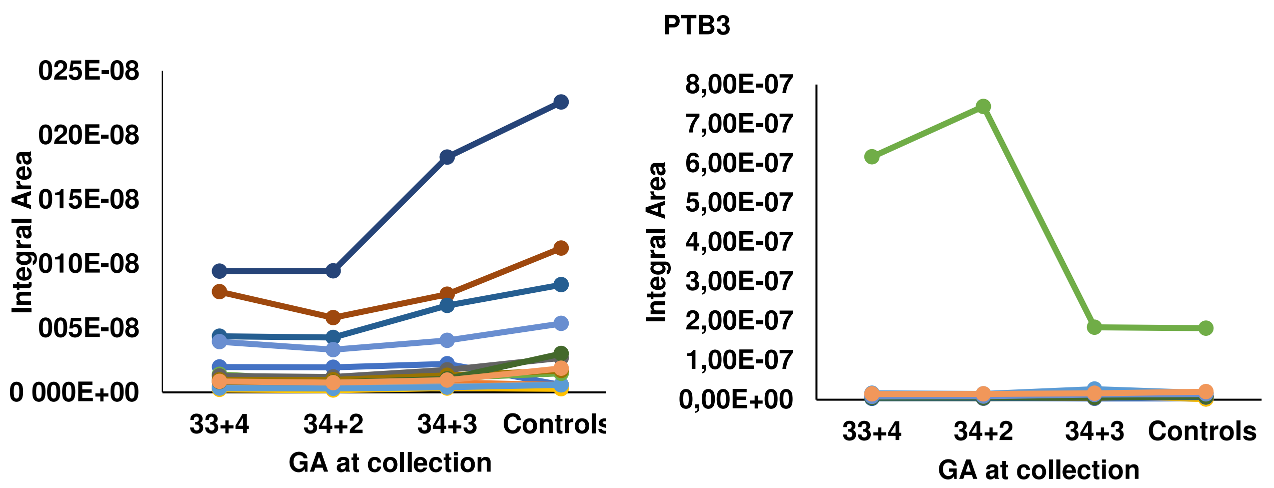
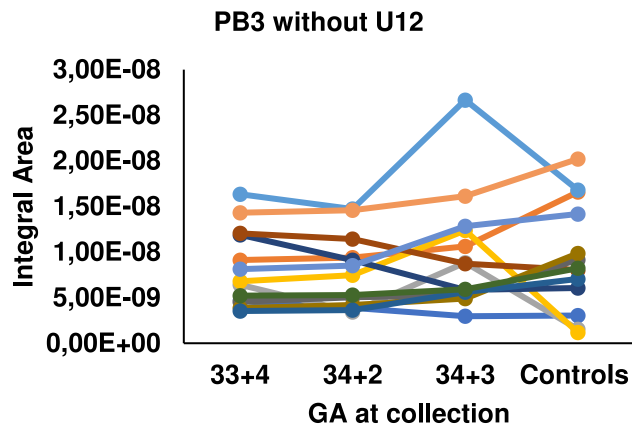
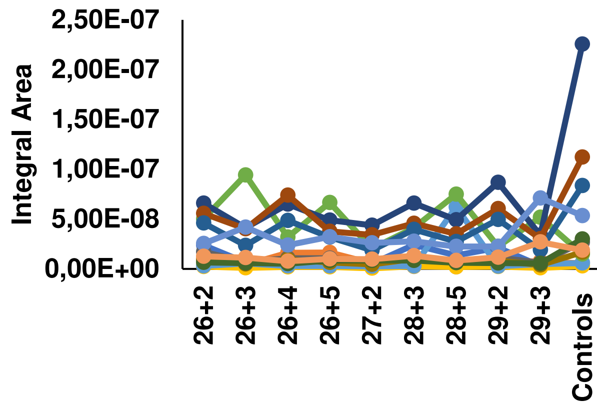


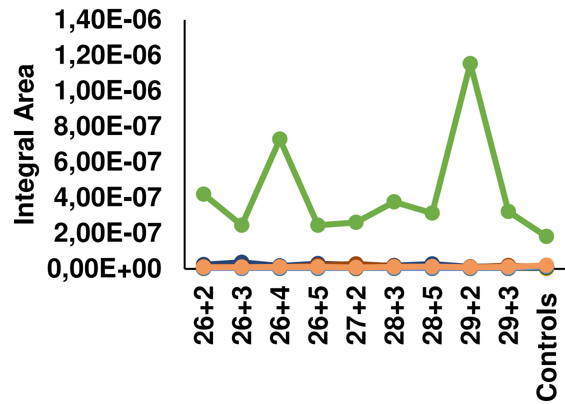
Figure 25 – Trajectory of a) metabolites and b) unassigned and regions of unassignements according gestational age at collection (**Closed to controls**).



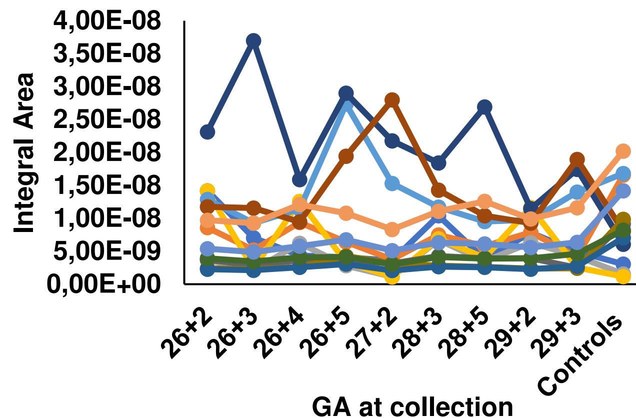


PTB5

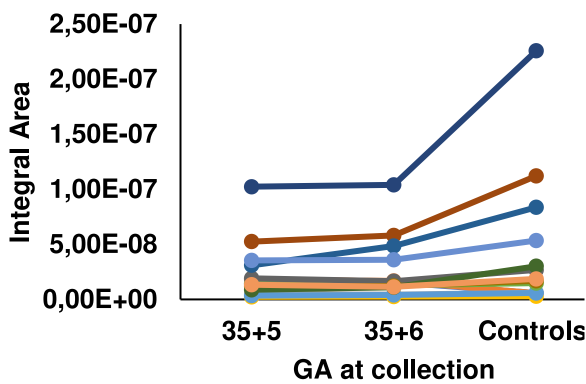


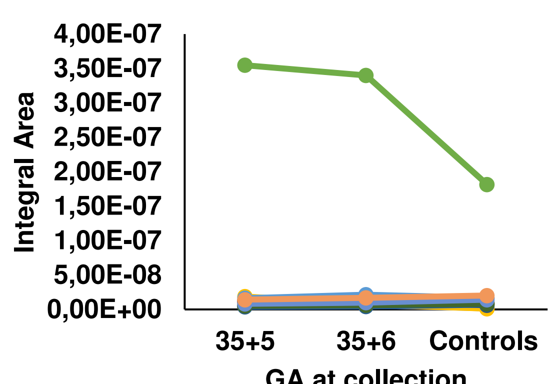


NB5 without U12

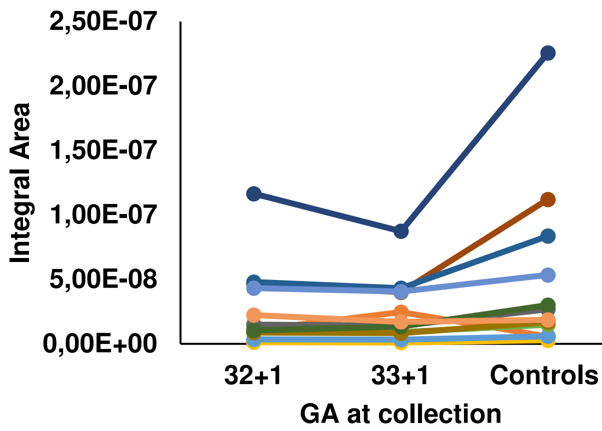


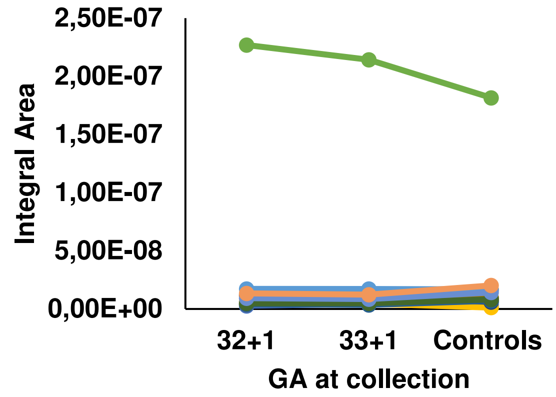
PTB6



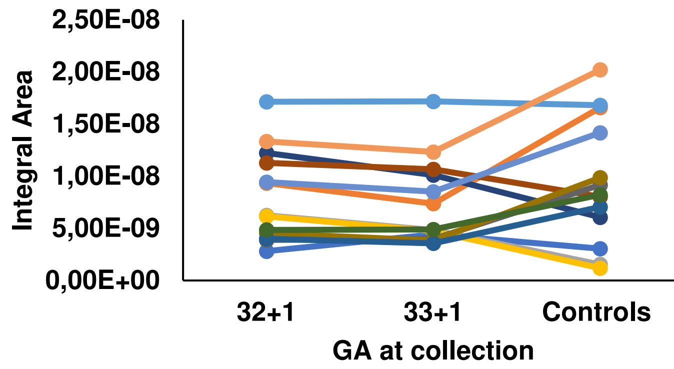


PTB9

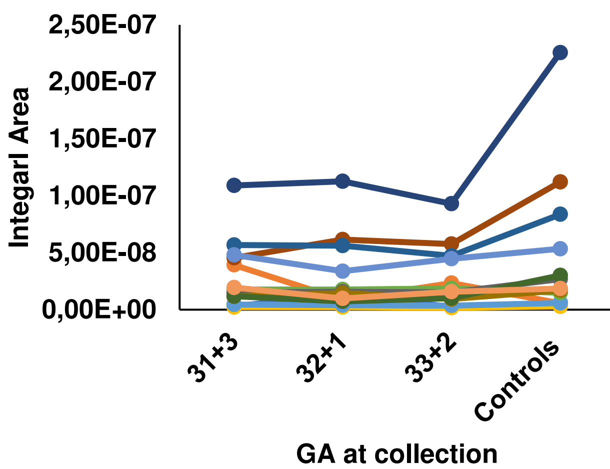


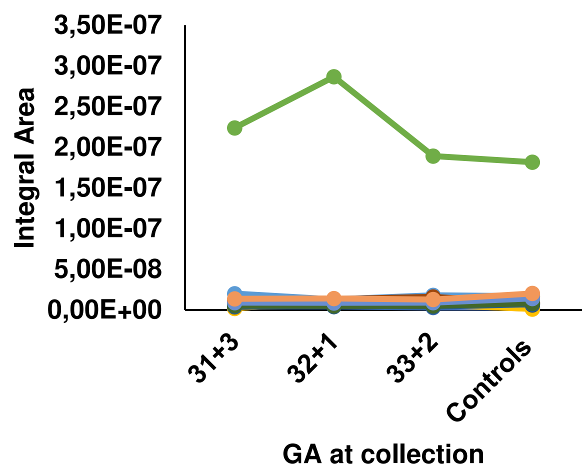


PTB9 without U12



PTB10





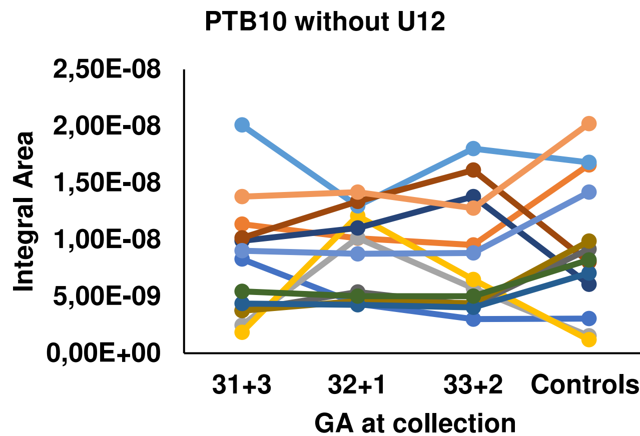


Figure 26 – Trajectory of a) metabolites and b) unassigned and regions of unassignements according gestational age at collection. **(Away from controls)**.

As referred above, PTB3, PTB5, PTB9 and PTB10 followed a visibly unstable trajectory compared with controls. In PTB3 there are a deviation in relation to controls and according to data, a decrease in ethanolamine and *N*-acetylneuraminic acid may be related with this alteration. Additionally, it is visible an increase of the unassigned signals 5, 6, 7 and 12. PTB9 and PTB10 are twins and both show a decrease in creatinine, DMA and U3, and an increase in formate. Additionally, PTB9 shows an increase in U2 and a decrease in U17, U18, Region1 and Region2. PTB10 shows an increase in U7, U13, U15 and a decrease in Region2. Is in PTB5 that are visible more metabolic alterations over time such as a decrease in ethanolamine, creatinine, DMA, *N*-acetylneuraminic acid and increase in glucose. In relation to unassigned metabolites is visible an increase followed by a decreased of U7, U13 and U15, and an increase of U12. An interesting observation was that when there is an increase in glucose the unassigned U12 decrease, so there may be a relationship between the two metabolites, which will be interesting to study in future.

According to graphs analysis and the corresponding table it is visible than some variations are common to those individuals: namely those in ethanolamine, creatinine, formate, DMA and *N*-acetylneuraminic acid,

5.2. Proposed metabolic interpretation of changes during premature newborns development until term time

According to *Table 9*, PTB5 has risk of infection, sepsis, hyponatremia, hyperbilirubinemia and a mother with 6 previous gestations has gestational diabetes. The conditions and his lower g.a. (EPT) may be responsible for an instable trajectory. Metabolic alterations over time are visible such as a decrease in ethanolamine, creatinine, DMA, *N*-acetylneuraminic acid and increase in glucose. In relation to unassigned metabolites is visible an increase followed by a decreased of U7, U13 and U15, and an increase of U12. Conditions such as Diabetes, Renal failure and nutrients such as proteins and fatty acids can altered the tryptophan pathway. Ethanolamine was found decreased suggesting an alteration on surfactants formations due to g.a.. Lower creatinine in PTB NB urine has been reported to immature tubular and vascular structures in the kidney. Compare with previous study on Chapter 4 none of the metabolites find significant relevant of EPT NB were found to vary in this study. PTB9 has transitory tachypnea and is twin of PTB10 without any conditions. These PTBs are VPT NBs and both show a decrease in creatinine, DMA and U3, and an increase in formate. DMA had been related with gut microflora and was found decreased. The reason for this variation is that the gut microflora of PTB NBs is poor. Additionally, PTB9 shows an increase in U2 and a decrease in U17, U18, Region1 and Region2. PTB10 shows an increase in U7, U13, U15 and a decrease in Region2. Compare with previous study on Chapter 4 formate and U2 were also found to vary in this study. PTB3 born due of membrane rupture, has tachypnea transitory, conjunctivitis and hyperbilirubinemia. Is a MLPT NB and shows a decrease in ethanolamine, *N*-acetylneuraminic acid and an increase of the unassigned peaks 5, 6, 7 and 12. Ethanolamine was found decreased suggesting an alteration on surfactants formations.All the metabolites found to vary in this PTB were also found significant on previously study on Chapter 4.

Chapter 6. Conclusion and future perspectives

This work employed a metabolomics approach to characterize PTB NBs urine composition using NMR spectroscopy.

In this study 49 metabolites were found in NB urine. The first study aimed at defining and characterizing urinary metabolic profile of NBs urine in prematurity. The results show a separation trend between PTB (in general and between the 3 groups of g.a.) and controls. Was found two outliers PTB17 (VPT) and PTB26 (MLPT) and to understand why, these two spectra were compared with those of controls and specific changes were found: increased formate, DMG, *m*-inositol, taurine (for both PTB17 and PTB26), galactose, citrate, succinate, acetone (for PTB26), NMND, histidine, 1-methyl-histidine, acetate (for PTB17); and decreased hippurate, creatinine, ethanolamine, DMA, betaine (for both PTB17 and PTB26), 4-HPA (for PTB26) and tyrosine (for PTB17). Compared with controls, 2 pairs of twin newborns were found to show significantly different urinary profiles. For each pair, one of the babies had a profile closer to controls (PTB5 on pair PTB17/PTB5 and PTB43 on pair RN26/RN43) and the other baby was significantly deviated from normality (PTB17 on pair PTB17/PTB5 and PTB26 on pair RN26/RN43). Specifically, for the PTB5/PTB17 pair, the deviant baby shown higher excretory levels of creatinine, citrate, succinate, acetone and lower of NMND, glucose and *m*-inositol. In addition, for the PTB26/PTB43 pair, the deviant individual showed higher levels of *m*-inositol, acetate, lactate, propylene glycol and lower of formate, succinate and acetone. Metabolism analysis through urine is not enough for observe directly the state of certain NBs, since several hypotheses (such as gender, delivery mode, birth weight, pathological conditions, weight at birth or medication) were excluded through spectra and clinical conditions analysis. Thus, clinic collaboration is required in accompaniment with metabolomic to improved results interpretation in order to obtain better conclusions. Outliers were removed and PLS-DA between PTB and controls shows that exists a g.a. dependence on the distance of each PTB NB to the controls. However not all NBs follow this trend and therefore, it has metabolic alterations that may be affecting its metabolic profile, moving away or approaching more than was correct (term time). Then, a comparation between different stages of prematurity was made, and results were found on formate, hypoxanthine and U2 (7,88 ppm) which may be related with lower stages of prematurity, while 2-Py, DMA, *N*-acetyl neuraminic acid, 4-DTA, U5 (5,10 ppm), U6 (5,08 ppm), U7 (4.33ppm), U12 (3,77 ppm), U13 (3,39 ppm), U15 (3,22 ppm), U17 (2.90ppm)

and U19 (2.76 ppm) may be related with later stages of prematurity. Since NMND, 1-methylhistidine, creatinine, ethanolamine, *cis*-aconitate, methylguanidine, 3-HBA, U3 (7.68 ppm), U16 (3.09 ppm), U18 (2.7 ppm), Region1 (2.78 -2.77 ppm) and Region2 (1.25-1.23 ppm) where found significant for all groups, may indicate that are relative with PTB.

In second study was made a analysis of PTB NB urinary trajectory in order to characterize the metabolic changes occurring during the NBs hospital stay. According to results analysis it is visible than some variations are common to those individuals: namely those in ethanolamine, creatinine, formate, DMA and *N*-acetylneuraminic acid.

The work presented in this thesis demonstrated that urine is a promising biofluid use on metabolomics for detection and prediction of disorders of the NB. It is important the use of larger cohorts for better and valid results. Infants are in constant development and the metabolic profiles of their urine changes constantly and significantly, especially in the first days of life. This is particularly evident for PTB NBs and during their development until they become apt to be released from hospital. The present work is one of the first longitudinal studies of very premature newborns, to the best of our knowledge. Although, a intensive study is needed in the future.

Bibliography

1. Sheiner E. The Long-Term Impact of Medical Complications in Pregnancy: A Window into Maternal and Fetal Future Health. CRC Press; 2017. Available from https://www.books.google.pt.
2. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. The Lancet. 2015 Jan;385(9966):430–40.
3. Blencowe H, Cousens S, Chou D, Oestergaard M, Say L, Moller A-B, et al. Born Too Soon: The global epidemiology of 15 million preterm births. Reprod Health. 2013 Nov;10(Suppl 1):S2.
4. Beck S, Wojdyla D, Say L, Betran AP, Merialdi M, Requejo JH, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ. 2010 Jan;88(1):31–8.
5. Portal do Instituto Nacional de Estatística [Web page]. [acessed in 2018 Set 19].
6. Goldenberg RL, Culhane JF, Iams JD, Romero R. Preterm Birth 1 Epidemiology and causes of preterm birth. The Lancet. 2008 Jan;371:10.
7. Greene MF, Creasy RK, Resnik R, Iams JD, Lockwood CJ, Moore T. Creasy and Resnik's Maternal-Fetal Medicine: Principles and Practice E-Book. Elsevier Health Sciences; 2009. Available from: https://pt.scribd.com.
8. Offiah I, ODonoghue K, Kenny L. Clinical Risk Factors for Preterm Birth. In: Morrison J, editor. Preterm Birth - Mother and Child [Ebook]. InTechOpen; 2012 [cited in 2018 Set 18]. Available from: http://www.intechopen.com.
9. Behrman RE, Butler AS. PRETEM BIRTH: causes, consequences and prevention. Washington, D.C. The National Academies Press. 2007.
10. Kliegman RM, Stanton BMD, Geme JS, Schor NF, Behrman RE. Nelson Textbook of Pediatrics E-Book. Elsevier Health Sciences; 2011. Available From: https://books.google.pt.
11. Perales JG, Pistelli IP, Junior A da SC. Doenças Respiratórias na Infância: Aspectos Biomoleculares, Clínicos e Cirúrgicos. Elsevier Brasil; 2016. Available from: https://books.google.pt.
12. Gallacher DJ, Hart K, Kotecha S. Common respiratory conditions of the newborn. Breathe. 2016 March;12(1):30–42.
13. Blackburn S. Brain Injury in Preterm Infants: Pathogenesis and Nursing Implications. Newborn Infant Nurs Rev. 2016 March;16(1):8–12.
14. Schwiertz A, Gruhl B, Löbnitz M, Michel P, Radke M, Blaut M. Development of the Intestinal Bacterial Composition in Hospitalized Preterm Infants in Comparison with Breast-Fed, Full-Term Infants. Pediatr Res. 2003 Sept;54(3):393–9.
15. Lewis ED, Richard C, Larsen BM, Field CJ. The Importance of Human Milk for Immunity in Preterm Infants. Clin Perinatol. 2017 March;44(1):23–47.
16. Canuto G, Costa JL, Cruz P, Souza A, Faccio A, Klassen A, et al. METABOLÔMICA: DEFINIÇÕES, ESTADO-DA-ARTE E APLICAÇÕES REPRESENTATIVAS. Quím Nova. 2017 Oct;41(1):75-91.
17. Alonso A, Marsal S, Julià A. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. Front Bioeng Biotechnol. 2015 March;3(23):1-20.
18. Cesare Marincola F, Corbu S, Lussu M, Noto A, Dessì A, Longo S, et al. Impact of Early Postnatal Nutrition on the NMR Urinary Metabolic Profile of Infant. J Proteome Res. 2016 Oct;15(10):3712–23.
19. Mussap M, Antonucci R, Noto A, Fanos V. The role of metabolomics in neonatal and pediatric laboratory medicine. Clin Chim Acta. 2013 Nov;426:127–38.
20. Fanos V, Iacovidou N, Puddu M, Ottonello G, Noto A, Atzori L. Metabolomics in neonatal life. Early Hum Dev. 2013 Jun;89:S7–10.
21. Vinayavekhin N, Homan EA, Saghatelian A. Exploring Disease through Metabolomics. ACS Chem Biol. 2010 Jan;5(1):91–103.
22. Rochfort S. Metabolomics Reviewed: A New "Omics" Platform Technology for Systems Biology and Implications for Natural Products Research. J Nat Prod. 2005 Dec;68(12):1813–20.
23. Bingol K, Brüschweiler R. Multidimensional Approaches to NMR-Based Metabolomics. Anal Chem. 2014 Jan;86(1):47–57.
24. Emwas A-H, Roy R, McKay RT, Ryan D, Brennan L, Tenori L, et al. Recommendations and Standardization of Biomarker Quantification Using NMR-Based Metabolomics with Particular Focus on Urinary Analysis. J Proteome Res. 2016 Fev;15(2):360–73.
25. Nagana Gowda GA, Raftery D. Recent Advances in NMR-Based Metabolomics. Anal Chem. 2017 Jan;89(1):490–510.

26. Jackman LM, Sternhell S. Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry: International Series in Organic Chemistry. Elsevier; 2013. Available from: https://books.google.pt.
27. Jacobsen NE. NMR Spectroscopy Explained: Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology. John Wiley & Sons; 2007. Available from: https://books.google.pt.
28. Wawer I, Diehl B. NMR Spectroscopy in Pharmaceutical Analysis. Elsevier; 2017. Available from: https://books.google.pt.
29. Keeler J. Understanding NMR Spectroscopy. John Wiley & Sons; 2011. Available from: https://books.google.pt.
30. Moolenaar SH, Engelke UFH, Wevers RA. Proton nuclear magnetic resonance spectroscopy of body fluids in the field of inborn errors of metabolism. Ann Clin Biochem. 2003 Jan;40(1):16–24.
31. Field LD, Li HL, Magill AM. Organic Structures from 2D NMR Spectra. John Wiley & Sons; 2015. Available from: https://books.google.pt.
32. Hoffmann E de, Stroobant V. Mass Spectrometry: Principles and Applications. John Wiley & Sons; 2013. Available from: https://books.google.pt.
33. Fanos V, Van den Anker J, Noto A, Mussap M, Atzori L. Metabolomics in neonatology: Fact or fiction? Semin Fetal Neonatal Med. 2013 Fev;18(1):3–12.
34. Bingol K. Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. High-Throughput. 2018 April;7(2):1-11.
35. Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013 April;9(2):280–99.
36. Mishur RJ, Rea SL. Applications of mass spectrometry to metabolomics and metabonomics: Detection of biomarkers of aging and of age-related diseases. Mass Spectrom Rev. 2012 Jan;31(1):70–95.
37. Noto A, Fanos V, Dessì A. Chapter Two - Metabolomics in Newborns. In: Makowski GS, editor. Advances in Clinical Chemistry [Ebook]. Elsevier; 2016 [acessed in 2018 Set 19]. p. 35–61Available from:http://www.sciencedirect.com.
38. Atzori L, Antonucci R, Barberini L, Locci E, Marincola FC, Scano P, et al. 1H NMR-based metabolomic analysis of urine from preterm and term neonates. Front Biosci Elite Ed. 2011 Jun;3:1005–12.
39. Trygg J, Holmes E, Lundstedt T. Chemometrics in Metabonomics. J Proteome Res. 2007 Fev;6(2):469–79.
40. Smolinska A, Blanchet L, Buydens LMC, Wijmenga SS. NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. Anal Chim Acta. 2012 Oct;750:82–97.
41. Biomarker Definitions Working group. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001 March;69(3):89–95.
42. Gil AM, Duarte D. Biofluid Metabolomics in Preterm Birth Research. Reprod Sci. 2018 Jul;25(7):967–77.
43. Ryan D, Robards K, Prenzler PD, Kendall M. Recent and potential developments in the analysis of urine: A review. Anal Chim Acta. 2011 Jan;684(1):17–29.
44. Mamas M, Dunn WB, Neyses L, Goodacre R. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. Arch Toxicol. 2011 Jan;85(1):5–17.
45. Mussap M, Noto A, Cibecchini F, Fanos V. The importance of biomarkers in neonatology. Semin Fetal Neonatal Med. 2013 Fev;18(1):56–64.
46. Luque de Castro MD, Priego-Capote F. The analytical process to search for metabolomics biomarkers. J Pharm Biomed Anal. 2018 Jan;147:341–9.
47. Varmuza K, Filzmoser P. Introduction to Multivariate Statistical Analysis in Chemometrics. CRC Press; 2016. Available from: http://books.google.pt.
48. Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, et al. Assessment of PLSDA cross validation. Metabolomics. 2008 March;4(1):81–9.
49. Joshi MS, Montgomery KA, Giannone PJ, Bauer JA, Hanna MH. Renal injury in neonates: use of “omics” for developing precision medicine in neonatology. Pediatr Res. 2017 Jan;81(1–2):271–6.
50. Jackson F, Georgakopoulou N, Kaluarachchi M, Kyriakides M, Andreas N, Przysiezna N, et al. Development of a Pipeline for Exploratory Metabolic Profiling of Infant Urine. J Proteome Res. 2016 Setp;15(9):3432–40.
51. Fanos V, Atzori L, Makarenko K, Melis GB, Ferrazzi E. Metabolomics Application in Maternal-Fetal Medicine. BioMed Research International. 2013 May;1-9.

52. Barker DJ. Fetal origins of coronary heart disease. BMJ. 1995 Jul;311(6998):171–4.
53. Trump S, Laudi S, Unruh N, Goelz R, Leibfritz D. 1H-NMR metabolic profiling of human neonatal urine. Magn Reson Mater Phys Biol Med. 2006 Nov;19(6):305.
54. Diaz SO, Pinto J, Barros AS, Morais E, Duarte D, Negrão F, et al. Newborn Urinary Metabolic Signatures of Prematurity and Other Disorders: A Case Control Study. J Proteome Res. 2016 Jan;15(1):311–25
55. Moltu SJ, Sachse D, Blakstad EW, Strømmen K, Nakstad B, Almaas AN, et al. Urinary Metabolite Profiles in Premature Infants Show Early Postnatal Metabolic Adaptation and Maturation. Nutrients. 2014 May;6(5):1913–30.
56. Brown JCC, Mills GA, Sadler PJ, Walker V. 1H NMR studies of urine from premature and sick babies. Magn Reson Med. 1989 Aug;11(2):193–201.
57. Foxall PJ, Bewley S, Neild GH, Rodeck CH, Nicholson JK. Analysis of fetal and neonatal urine using proton nuclear magnetic resonance spectroscopy. Arch Dis Child Fetal Neonatal Ed. 1995 Nov;73(3):F153–7.
58. 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 Fev;75(2):367–73.
59. 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 Fev;15:203–9.
60. 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 April;214(4):513.e1-513.e9.
61. Liu J, Chen X-X, Li X-W, Fu W, Zhang W-Q. Metabolomic Research on Newborn Infants With Intrauterine Growth Restriction. Medicine (Baltimore). 2016 April;95(17).
62. Jelliffe-Pawlowski LL, Norton ME, Baer RJ, Santos N, Rutherford GW. Gestational dating by metabolic profile at birth: a California cohort study. Am J Obstet Gynecol. 2016 April;214(4):511.e1-511.e13.
63. Moco S, Collino S, Rezzi S, Martin F-PJ. Metabolomics perspectives in pediatric research. Pediatr Res. 2013 Jan;73(4–2):570–6.
64. Longini M, Giglio S, Perrone S, Vivi A, Tassini M, Fanos V, et al. Proton nuclear magnetic resonance spectroscopy of urine samples in preterm asphyctic newborn: A metabolomic approach. Clin Chim Acta. 2015 April;444:250–6.
65. Fanos V, Pintus MC, Lussu M, Atzori L, Noto A, Stronati M, et al. Urinary metabolomics of bronchopulmonary dysplasia (BPD): preliminary data at birth suggest it is a congenital disease. J Matern Fetal Neonatal Med. 2014 Oct;27(sup2):39–45.
66. Berben L., Sereika M. S., Engberg S. Effect size estimation: Methods and examples.International Journal of Nursing studies. 2011 January 26;1039-1047.