METABOLOMICS AND NEONATAL SEPSIS:
A NEW APPROACH FOR EARLY BIOMARKERS DISCOVERY

Coordinatore: Ch.mo Prof. Carlo Giaquinto
Supervisore: Ch.mo Prof. Eugenio Baraldi

Dottorando: Veronica Mardegan
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>3</td>
</tr>
<tr>
<td>Aims</td>
<td>16</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>17</td>
</tr>
<tr>
<td>Results</td>
<td>22</td>
</tr>
<tr>
<td>Discussion</td>
<td>27</td>
</tr>
<tr>
<td>Conclusions</td>
<td>31</td>
</tr>
<tr>
<td>References</td>
<td>32</td>
</tr>
</tbody>
</table>
ABSTRACT

Background: Neonatal sepsis is a complex infection-induced systemic inflammatory response syndrome and it is a main cause of mortality and neurologic sequelae in newborns. An early and accurate detection of sepsis is mandatory in neonates, since the clinical course of the infectious process can be fulminant, leading to septic shock and death within hours after the first clinical symptoms. The gold standard for diagnosis of neonatal sepsis is blood culture, but it is time-consuming and false negative results are not rare. To date, there is still no reliable biochemical marker of neonatal sepsis.

Aim of the study: To compare the metabolic profile of urine collected within 72 hours of birth between preterm neonates affected by early onset sepsis (EOS) and healthy preterm infants, searching for a specific metabolite or metabolic profile enabling the early identification of preterm newborns prone to develop EOS.

Materials and Methods: Each preterm neonate admitted to the Neonatal Intensive Care Unit was eligible for recruitment. Infants who developed a septic episode within 72 hours of birth were enrolled as cases. Infants who did not develop a septic episode were enrolled as controls. For each subject, a urine sample was collected within 72 hours of birth. The urine samples underwent untargeted metabolomic analysis using mass spectrometry combined with ultra-performance liquid chromatography. The data obtained were analyzed using multivariate and univariate statistical data analysis tools.

Results: One-hundred and twenty-three subjects were enrolled in the study. Seventeenth neonates were affected by EOS. Seventeenth gestational age-matched newborns were enrolled as controls. Metabolomic untargeted analysis on urine samples collected within 24 hours of birth revealed an evident clustering of subjects (septic versus non-septic neonates). The diagnostic performance of urine metabolome resulted comparable to those of PCR, as documented by ROC curves.

Conclusions: Neonates with EOS showed a specific metabolic profile compared to those of newborns not affected by sepsis at the onset of infection, allowing their clear discrimination with
the use of untargeted metabolomics analysis. Results of this research support the effectiveness of metabolomics in exploring biochemical pathways of neonatal sepsis, potentially providing novel putative biomarkers for early diagnosis, guide to antibiotic therapy, and monitoring of disease progression. Future perspectives will be addressed to introduce these novel metabolomics biomarkers in clinical routine practice, ensuring short laboratory turnaround time and 24 hours bedside availability.
BACKGROUND

Neonatal mortality and main neonatal morbidities

During the last decades, great improvements were done in perinatal and neonatal care. Treatment of pregnant women expected to delivery preterm with antenatal steroids to induce fetal lung maturation, intrapartum antibiotic therapy of women with positive rectal-vaginal swab, introduction of new therapies, for example endotracheal surfactant to enhance lung compliance, and new treatment strategies such as “gentle ventilation”, significantly affected neonatal outcome. As a consequence, neonatal mortality has continuously decreased worldwide, and even death related to preterm birth declined.\textsuperscript{1-6} Italian newborn mortality is among the lowest worldwide, with a mortality rate of 2.1 per 1000 live births.\textsuperscript{1}

While the improvement in neonatal survival was substantial in the last decades of the 19\textsuperscript{th} century, there has been a small change in developed countries over the last years.\textsuperscript{1,2} Since the main causes of newborn deaths are prematurity and low-birth-weight, infections, asphyxia, birth trauma and congenital abnormalities (accounting for nearly 80\% of deaths in this age group), strategies to prevent and effectively treat these conditions are warranted to achieve a further decline in mortality.\textsuperscript{7-9}

An equally relevant goal in neonatal research is the minimization of long-term morbidities secondary to neonatal injuries. Most of premature babies still survive after experiencing at least one major neonatal morbidity known to be associated with both short- and long-term adverse consequences and despite the progress in neonatal survival, the burden of neonatal morbidities related to preterm birth, such as bronchopulmonary dysplasia (BPD), white matter injury (WMI), necrotizing enterocolitis (NEC) and retinopathy of prematurity (ROP), remains important. Infants weighing 501 to 750 g had the greatest decrease in mortality but the least change in survival with major morbidity.\textsuperscript{6} Neurological outcome is the most relevant concern and preterm-birth complications, asphyxia and sepsis are documented as the most important preventable reasons of
neurological sequelae worldwide. Therefore great efforts should be done to predict and diminish their occurrence and to optimize therapeutic strategies. Particularly, sepsis frequently affects premature babies and it is supposed to play a key role in most inflammatory disorders that cause or enhance main preterm morbidities (BPD, WMI, NEC and ROP).\textsuperscript{10}

Since we have recently demonstrated that early metabolic dysregulations may play a key role in the pathogenesis of some relevant neonatal conditions, such as preterm birth and bronchopulmonary dysplasia\textsuperscript{11}, it seems particularly promising to apply the metabolomic approach for a better understanding of neonatal sepsis.

**Neonatal Sepsis**

Neonatal sepsis is an infection-induced systemic inflammatory response syndrome and it is a main cause of mortality and morbidity in newborns.\textsuperscript{3,8,9,12}

In high-income countries, about 1-2\% of all live births are affected by neonatal infections and a recent meta-analysis estimated an incidence of 3·0 million cases of sepsis in neonates.\textsuperscript{13,14}

According to the onset of age, sepsis is typically classified as early-onset neonatal sepsis (EOS, infection occurring within 3 days after birth), and late-onset neonatal sepsis (LOS, infection occurring at > 3 days of life).\textsuperscript{15,16}

EOS reflects transplacental or, more frequently, ascending infections from the maternal genital tract, and symptoms are frequently severe. The most commonly reported pathogens are Group B Streptococcus (GBS, 37.8\%), E. coli (24.2\%), viridans Streptococci (17.9\%), Staphylococcus aureus (4.0\%), and Haemophilus influenzae (4.0\%).\textsuperscript{17-19} The annual incidence of EOS has been estimated around 0.5-1.2 per 1,000 newborns, and it significantly increases with the lowering of gestational age. Among term infants, the leading infection is GBS (0.22 cases per 1,000 live births), whilst among preterm infants, E. coli is the most common infection (1.18-10.4 cases per 1,000 live births). Case fatality also is inversely related to gestational age and birth weight, reaching the
highest values (50-60%) in ELBWI (extremely low birth weight infants, birth weight <1000 g) born at 22-24 weeks gestation.\textsuperscript{20} Mortality is also strongly related to aetiology, ranging from 35.3% for E. coli infections to 2.5% for infants with viridans Streptococci.\textsuperscript{17-19,21-23}

A really effective intervention for the reduction of EOS incidence was the introduction since the early 2000s of the antenatal screening for GBS colonization and intrapartum antimicrobial prophylaxis to colonized women. The incidence of invasive early-onset GBS disease decreased by more than 80%, from 1.8 cases/1000 live births in the early 1990s to 0.26 cases/1000 live births in 2010. Intrapartum prophylaxis was similarly effective among term and preterm infants, but it hasn’t impact the incidence of LOS.\textsuperscript{24,25}

Beyond GBS colonization, other risk factors of EOS were identified, such as premature rupture of membranes (PPROM), maternal fever and chorioamnionitis, meconium stained amniotic fluid, but their role is still matter of investigation.\textsuperscript{26,27}

LOS is rare in term babies, whilst it represents a frequent complication of premature birth. This condition is due to the intrinsic immaturity of the immune system and to predisposing issues associated to Neonatal Intensive Care Unit (NICU) hospitalisation, such as patent ductus arteriosus, necrotizing enterocolitis, mechanical ventilation, bronchopulmonary dysplasia, the failure of early enteral feeding with breast milk, administration of total parental nutrition, use of central venous catheters, surgery and underlying respiratory and cardiovascular diseases.\textsuperscript{16,28,29}

Whilst epidemiological studies have observed a global reduction in EOS during the last decades, probably due to the use of prophylactic intrapartum antibiotics to prevent infections caused by GBS, the incidence of LOS has increased in parallel with the improved survival of premature infants, indicating the role of hospitalisation in the pathogenesis of neonatal LOS.\textsuperscript{16} The incidence of LOS in the NICU is inversely related to birth weight and gestational age and it ranges from 25–36% in very low birth weight infants (VLBWI: birth weight ≤1500 g) to 6.2–10% in late-preterm (GA: 34–37 weeks) infants. It was estimated that one in three VLBWI and one in two ELBWIs will develop at
least one episode of LOS during the hospital stay, thus leading to a huge increase of mortality (18-21% in VLBW and 20-40% in ELBW) and hospitalisation length in this population. Gram-positive organisms are the most common pathogens in LOS, with coagulase-negative staphylococci accounting for 40-78% of infections. Other bacteria commonly involved in LOS are Staphylococcus aureus (7,8-20%), Klebsiella (4-12%), E. coli (4,9-7,9%), Enterobacter spp., Pseudomonas spp. and Candida (5,5-10%). Neonates with gram-negative sepsis, fungemia or Pseudomonas LOS have significantly higher risk of overall mortality, compared with those with gram-positive LOS.\textsuperscript{16,17,28-32}

But, despite a significantly lower mortality rate, CONS sepsis is associated with a risk of neurodevelopment sequelae, such as cognitive and psychomotor impairment, cerebral palsy and vision impairment, equal to Gram-negative bacteria and fungi, indicating that CONS are capable to exert a long-term detrimental effect on the host, particularly on ELBW infants.\textsuperscript{16,33}

Although most neonatal sepsis are primary bacteremia, 17.6% of LOS are associated with a concurrent infectious focus, including meningitis (5.4% of LOS), pneumonia (3.8%) and NEC (2.2%) and this association correlates with significantly higher rates of infectious complication.\textsuperscript{34}

Particularly worrisome is the central nervous system (CNS) concomitant infection, as it highly enhances the risk of neurological sequelae. The incomplete development of the neonatal blood–brain barrier makes this happening not unlikely, exposing newborns, particularly if preterm, to a great threat to their health and life. Recent studies suggest that bacteraemia could trigger cerebral injury even without penetration of viable bacteria into the CNS, and sepsis represents by itself a risk factor for neurodevelopmental impairments, including cerebral palsy, low psychomotor developmental index and deafness. VLBW infants with sepsis in the neonatal period have poor long-term neurodevelopment outcome compared with those without sepsis. White matter injury, particularly periventricular leukomalacia, plays a key role in the correlation between sepsis and neurodevelopmental damage. It seems to be induced by a multifactorial process involving the production of pro-inflammatory cytokines, increased blood–brain barrier permeability and hypoxic ischemic events resulting from hypotension and impaired autoregulation of cerebral blood flow.
Mechanisms of damage are mediated by loss of immature preoligodendrocytes, which are particularly susceptible to oxidative stress, inhibition of neuronal precursor cell proliferation and activation of astrogliosis.\textsuperscript{14,32,33,35,36}

**Diagnosis of neonatal sepsis**

Taking into account this background, it comes to light the need for an accurate and early detection of neonatal sepsis, so that a broad-spectrum antibiotic therapy can be promptly started in affected newborns to stop the fulminant clinical course that could lead to septic shock and death within hours after the first clinical symptoms. On the other hand, it is also very important to rule out this diagnosis in non-infected patients, to avoid unnecessary treatment that might expose babies to prolonged hospitalisation and potential adverse events. This could also reduce healthcare costs and limit the worrying increase of antibiotics resistance. The uncertainty of sepsis diagnosis, in fact, commonly results in unnecessary and prolonged empiric antibiotic treatment and this policy lead to an alarmingly high degree of antibiotic resistance worldwide.\textsuperscript{37-41}

Accurate diagnosis of neonatal sepsis is challenging. Blood culture is still considered the gold standard. However, results are not quickly available and false negative responses are not rare, since bacteraemia is often intermittent and the blood sample is frequently small in volume in preterm neonates. Moreover, intrapartum antibiotic treatment may further reduce the diagnostic value of blood culture.\textsuperscript{17,19} Culture-confirmed bacterial sepsis is consequently quite uncommon and typically represents no more than 5-50% of all clinically-suspected neonatal infections.\textsuperscript{19,40}

Recently, molecular-based methods have emerged as promising diagnostic tools. Polymerase Chain Reaction, a technology based on the extraction of microbial DNA from biological samples and the subsequent sequencing or hybridisation of species-specific gene regions, was widely investigated for the detection of micro-organisms and it showed a high sensibility.\textsuperscript{16} Nevertheless this cost consuming technique is not widely diffused, and it does not allow the detection of antibiotic-resistance profile.
Neonatal sepsis is therefore mainly suspected based on clinical presentation, in the presence of non-specific signs and symptoms (respiratory distress, apnoea, tachypnea, tachycardia, fever, hypothermia, lethargy, hypotonia, feeding intolerance, poor perfusion, irregular cardiac rhythms, metabolic acidosis) that could be manifestations of many other clinical conditions, such as transient tachypnea of newborns, respiratory distress syndrome, patent ductus arteriosus, congenital heart disease and so on.\textsuperscript{15,42,43} It should be also considered that the disease may be present without the appearance of clinical symptoms and 20\% of infected term infants seem healthy and blood culture is performed in these babies only for maternal risk factors.\textsuperscript{17,22}

For this reason, biochemical markers could help in improving diagnostic accuracy. Most widely used biomarkers are white blood cells count, neutrophil count, immature to total neutrophil count ratio, platelet count, C reactive protein and procalcitonin, but none of them can be considered an absolutely reliable sepsis indicator in newborns.

White blood cell (WBC) count, absolute neutrophil count and the ratio of immature to total neutrophils are non-specific for the diagnosis of sepsis. Total WBC count give a poor positive indicative value for sepsis, neutrophil values depend upon the individual's age and neutropenia can be a sensitive measure but depends upon delivery method, gestational age and altitude. Platelet count is used as early diagnostic testing for neonatal sepsis, but this test is not very specific nor helps in monitoring for treatment procedure.\textsuperscript{17}

CRP is the most widely studied and used biomarker for the detection of neonatal sepsis. It is an acute phase protein produced by the liver. CRP levels take 6-12 hours to alter after the start of infection, and it has 24-48 hours of half-life. This kinetic explains the low sensitivity of CRP for the early detection of sepsis. It shows a specificity of 94.8\% and sensitivity of 67.1\%. Serial CRP assessment helps in monitoring the response to treatment and the negativity of CRP levels indicate the end of the infective process, hence implying antibiotic therapy discontinuation. Nevertheless, CRP as sepsis biomarker has lots of limitations. It has a physiologic rise after birth and it elevates in some non-infectious conditions, such as inflammation, traumatic or ischemic tissue injuries,
meconium aspiration syndrome and hemolysis, thus limiting its potential for diagnosis of infection. It also seems to be influenced by some antenatal and perinatal conditions, such as maternal pregnancy-induced hypertension, premature rupture of membranes, duration of active labor, prenatal steroids and intrapartum antimicrobial prophylaxis. Furthermore, preterm infants generally exhibit small rise in CRP levels upon infection and effects of birth weight and gestational age on kinetics of CRP has not been well investigated. Finally, valid age specific reference values are not available, whilst many physiological and metabolic processes differ significantly from birth to later time periods.

Procalcitonin is a prohormone of calcitonin. It is an acute phase reactant, and is responsible for vascular response and immunomodulation related to systemic inflammatory response syndrome. PCT is synthesized by hepatocytes and monocytes within 2-4 hours of bacterial infection, then elevates in next 6-8 hours and its level stays for the next 24 hours. PCT has 24-30 hours half-life. It seems to show sensitivity and specificity values respectively of 85% and 54%, but reports in the literature are contradictory. Recently Stocker et al. demonstrated that PCT is a useful tool in reducing antibiotic therapy in neonates with suspected EOS. Optimal cutoff ranges in term and preterm newborns are still matter of debate, since values useful in adults and older children are too low for babies with a more reactive immune system. PCT levels physiologically rise at birth, up to 10 ng/ml within the first 24 postnatal hours, presumably in relation to intestinal bacterial colonization, and this limits notably its usefulness in EOS diagnosis. It seems that PCT physiologic rise in healthy preterm neonates is earlier, higher, and longer than in healthy term babies and this reasonably could happen in infected patients too. PCT falsely rise in non-infectious conditions like respiratory distress syndrome, hemodynamic failure, birth asphyxia, neonatal hypoxemia and intracranial hemorrhage and in relation to perinatal factors, such as chorioamnionitis and maternal pre-eclampsia; this could limit considerably its specificity.
A variety of other sepsis biomarkers (serum amyloid A, lipopolysaccharide binding protein, hepcidin, cytokines, presepsin, fibronectin, neopterin…) have been proposed in the last years, but most of them are still matter of study and clinical feasibility hasn’t been demonstrated yet. Therefore, no currently available test is able to provide perfect diagnostic accuracy, and false-negative as well as false-positive results may occur. The combination of multiple biomarkers may enable a fast and accurate diagnosis of neonatal sepsis, but to date the optimal selection of analytes has not been determined.

**Our setting**

Almost 350-400 newborns per year are admitted to the Neonatal Intensive Care Unit of the Women’s and Children’s Health Department of Padua. Among these patients, about 110 are Very Low Birth Weight Infants (birth weight <1500 g) and about 50 are Extremely Low Birth Weight Infants (birth weight < 1000 g).

Approximately 40% of admitted patients, mostly preterm infants, develop at least one event of suspected or proven sepsis during hospitalization.

**Metabolomics**

Metabolomics is the most recent of the “omic” sciences and it represents the analysis and interpretation of global metabolic data of a complex biological system. These metabolic data result from the information enclosed in the genetic code as well as from the effect of physiological factors (eg, diet, age, commensal microorganisms), pathological conditions, environmental agents, and exposure to drugs or toxic agents. Metabolomics is therefore considered the “omic” science that comes closest to phenotype expression.

Two approaches can be applied in metabolomics: in targeted metabolomics, a fixed subset of certain metabolites (generally up to 500) is studied, and this method is generally used to verify a
specific hypothesis; in untargeted profiling, all metabolites in a sample are measured (even with some detected metabolites not yet identified), and this approach is preferred for hypothesis-generating studies.55

As a matter of fact, the study of metabolites of a biological system can be managed in different methods. Metabolite targeting aims to identify a specific metabolite. Metabolite profiling allows the identification and quantification of a particular set of metabolites belonging to a specific metabolic pathway. Metabolite fingerprinting tries to detect the metabolic characteristics that discriminate between groups of subjects, without any a priori hypothesis. It does not necessarily involve identifying each metabolite, but its goal is the discrimination of metabolite patterns associated with a given pathological condition. Since it is not guided by a researcher’s hypotheses, it is open to new findings and unexpected or even unknown metabolites may turn out to be important in characterizing specific groups of subjects, so new pathophysiological hypotheses may be formulated.

Metabolomics studies are typically performed on biofluids, by analysing them with platforms such as nuclear magnetic resonance (NMR) spectroscopy or mass spectrometry (MS). $^1$H-NMR spectroscopy generates different signals for each metabolite in the NMR spectrum, thus favoring the identification of proton-containing metabolites. The sample, positioned in a NMR glass tube in a magnetic field, is excited with a radio frequency pulse; alternation between the lower and higher energy spin states of the electrons generates a resonance which is unique for every substance, depending on its chemical structure. Results are showed as peaks across a spectrum, and the area under each peak represents the relative concentration of that specific metabolite; by comparing the relative concentration to an appropriate reference signal, a precise quantification of the metabolite is provided. Advantages of $^1$H-NMR are minimal sample preparation, non-destruction of biofluids analyzed and execution speed.55

MS is a powerful method for identifying and quantifying metabolites, and it generates a spectrum in which biocompounds are represented according to their masses (m/z). Liquid chromatography or
gas chromatography, which improve mass separation, often precede MS. It is generally considered more sensitive than $^1$H-NMR, but MS-based platforms are more laborious and more destructive for the sample. Both techniques can be performed on very small samples (generally minimum 20 µL). Datasets deriving from NMR and MS analysis can be interpreted using appropriate multivariate statistical approaches (the “pattern recognition methods”). These approaches can be classified in unsupervised and supervised. The first ones, such as principal component analysis (PCA), represent data by means of plots that the human eye can interpret as different cluster for distinctive subjects’ categories (eg, healthy vs ill patients). The supervised methods, such as partial least squares - discriminant analysis (PLS-DA), use a training set of classified samples to create a mathematical model that is then used to test an independent dataset. This method enables us to predict which group a new sample belongs to according to its spectra characteristics.

From a clinical standpoint, metabolomics has the potential to contribute to the characterization of specific biomarkers of early diagnosis, prognosis or prediction of therapy response. For this reason, the metabolomics approach is considered promising with the view of the development of a more personalised medicine.56

**Metabolomics and sepsis**

Specifically in sepsis, metabolomics is a very promising field of research. Sepsis is in fact the result of the complex interactions between the host and the insulting pathogen. Sepsis-induced alterations in the genome, transcriptome and proteome are reflected in changes in the concentration of the small molecules in biological fluids and tissues. Since the impact of infectious perturbation on the metabolome results in changes that are highly dynamic and occur over short periods of time, metabolomics is an approach more suitable than the “static” genomics, transcriptomics and proteomics to study sepsis phenomenon. Metabolomics could therefore shed light on the complex biological pathways underlying the pathophysiology of sepsis.55
Recent published researches in adult population showed that metabolomics can discriminate between septic and non-septic patients and that metabolic profile markedly differs between healthy patients and patients with systemic inflammatory response syndrome or sepsis.\textsuperscript{57,58} A logistic regression model based on 6-metabolites performed better than conventional laboratory and clinical parameters in prediction of bacteremic sepsis.\textsuperscript{59}

Metabolomics can also forecast the severity of sepsis and septic shock in adult patients, accurately predicting mortality. In a study by Ferrario et al., early changes in plasma levels of some glycerophospholipids correlated with long-term survival of subjects.\textsuperscript{60} Liu et al. demonstrated that metabolites in the tricarboxylic acid cycle, amino acids, and energy metabolism-related metabolites were up-regulated in serum of non-survivors, whereas those in the urea cycle and fatty acids were generally down-regulated.\textsuperscript{61} Garcia-Simon et al. showed that that NMR metabolic profiling might be helpful for determining the metabolomic phenotype of worst-prognosis septic adult patients in an early stage; negative prognosis patients presented higher values of ethanol, glucose and hippurate, and on the contrary, lower levels of methionine, glutamine, arginine and phenylalanine.\textsuperscript{62} In a study by Langley et al., metabolites of the fatty-acid transport and $\beta$-oxidation, gluconeogenesis, and the citric acid cycle differed consistently among patients who would ultimately die from those who would survive, and they diverged more as death approached.\textsuperscript{63} The global metabolomic profile in plasma broadly differed between survivors and non-survivors of community acquired pneumonia and sepsis in a study by Seymour et al.; broad differences were present in pathways of oxidative stress, bile acid metabolism, stress response and nucleic acid metabolism.\textsuperscript{64}

An interesting further application of metabolomics is the discrimination of different triggering factors. By means of lipid metabolites analysis, Neugebauer et al. demonstrated that it is possible to distinguish patients with community-acquired pneumonia (CAP) and sepsis as compared to patients with sepsis due to other infections.\textsuperscript{65} To et al. showed that 13 metabolites can distinguish CAP affected patients from the non-infected ones, and from those who had an extrapulmonary infection or those who had a non-CAP respiratory tract infection.\textsuperscript{66}
Studies investigating the role of metabolomics in the diagnostic work-up of paediatric and neonatal sepsis are limited. As described by Mickiewics and others, through changes in serum metabolic profile detected with NMR spectrometry, it is possible to distinguish between children with septic shock, children with non-infectious systemic inflammatory response syndrome and healthy infants. In particular, babies with septic shock have an increased serum concentration of six metabolites when compared with healthy newborns. In addition to these potential diagnostic applications, the metabolomic approach turned out to be more accurate in determining the mortality risk in septic patients in comparison with PCT levels and PRISM III-APS score (a score commonly used in Intensive Care Units for the definition of mortality risk).\textsuperscript{67} The same author demonstrated that combining metabolomic and inflammatory protein mediator profiling early after presentation may differentiate children with sepsis requiring care in a Pediatric Intensive Care Unit (PICU) from children with or without sepsis safely cared for outside a PICU.\textsuperscript{68} In a case report with two control groups, Ambroggio et al. showed that the urine concentration of metabolites previously associated with sepsis (e.g., 3-hydroxybutyrate, carnitine, and creatinine) were higher in a patient with fatal MRSA (methicillin-resistant Staphylococcus aureus) pneumonia compared with those of patients with influenza pneumonia and healthy controls. Interestingly, these metabolic changes in the urine preceded the clinical severe sepsis phenotype, suggesting that detection of the extent of metabolic disruption can aid in the early identification of a sepsis phenotype in advance of the clinical diagnosis.\textsuperscript{69}

Metabolomics is gaining importance also in neonatology. Many studies analyzed different biofluids (serum, urine, cord blood, saliva, amniotic fluid, breast milk) with NMR o MS platforms, with particular interest in prematurity, intrauterine growth retardation, inborn errors of metabolism, perinatal asphyxia, sepsis, necrotizing enterocolitis, kidney disease, bronchopulmonary dysplasia, and cardiac malformation and dysfunction.\textsuperscript{70}

Regarding neonatal sepsis, only a few studies have been conducted to date.
Fanos and colleagues applied metabolomics to the analysis of urine of 25 neonates, showing that a specific urine metabolic profile can be detected in septic neonates: nine infected newborns with early- and late-onset sepsis were distinguished from 16 controls after urine analysis by $^1$H-NMR and GC-MS. The molecules responsible for the differences in the metabolic profiles were glucose, lactate and acetate, the urine content of which was increased in septic neonates compared to controls, while 2,3,4-trihydroxybutyric acid (THBA), ribitol, ribonic acid and citrate concentration was decreased.\(^{71}\)

A recent study by Sarafidis and others revealed that neonates with confirmed and possible late onset sepsis at the onset of clinical manifestations showed a different metabolic profile compared to those without sepsis, allowing their clear discrimination with the use of $^1$H-NMR and LC-MS/MS-based urine analysis. Neonates with confirmed and possible sepsis exhibited similar metabolic profiles under the applied methods, thus indicating similar underlying biological derangements and confirming the validity of clinical practice to treat these patients in the same way of newborns who have a positive blood culture. It is worth noting that no discrimination between septic and non-septic neonates could be made on day 10 of sepsis, suggesting that metabolomics could be a reliable tool to monitor response to therapy.\(^{72}\)

Recently, Fattuoni et al. demonstrated that urinary metabolomics can discriminate preterm neonates born to mothers with and without histological chorioamnionitis. Glutamate metabolism, mitochondrial electron transport chain, citric acid cycle, galactose metabolism, and fructose and mannose degradation metabolism were the most significantly altered pathways.\(^{73}\)

Only one study compared urinary metabolomic profile of a preterm patient affected by invasive fungal infection and those of healthy controls; the profile of the septic newborn was unique and, interestingly, it was possible to monitor the efficacy of therapy in improving patient health.\(^{74}\)

Although these are preliminary results which need to be confirmed in larger studies, they are certainly promising.
AIMS OF THE PROJECT

Primary aim:
The main aim is to compare the metabolic profile of urine collected within 72 hours of birth between preterm neonates affected by early onset sepsis and healthy preterm infants.

Secondary aim:
The search for a specific metabolite or metabolic profile enabling the early identification of preterm newborns prone to develop early onset sepsis.
MATERIALS AND METHODS

Ethical approval for the study was obtained from the local ethics committee and informed consent was collected from parents.

**Inclusion criteria**

Each infant had to meet all the following criteria to be enrolled in the study: admission to the Neonatal Intensive Care Unit, gestational age <37 weeks’ gestation, written informed consent for participation of a legally acceptable representative.

Babies who developed a septic episode within 72 hours of birth (EOS) were enrolled as *cases*.

Babies who did not develop a septic episode (neither EOS nor LOS) were enrolled as *controls* (aged matched control group).

**Definition of neonatal sepsis:**

Presumed sepsis as defined by presence of any three clinical or laboratory criteria from the list below OR

Confirmed bacterial sepsis as defined by positive blood and/or liquor culture and at least one clinical or one laboratory criterion from the list below:

- **Clinical criteria:** hyperthermia or hypothermia (≥ 38°C or ≤ 36°C), hypotension or impaired peripheral perfusion or mottled skin, apnoea or increased oxygen requirement or increased requirement for ventilatory support, bradycardic episodes or tachycardia, worsening feeding intolerance or abdominal distension, lethargy or hypotonia or irritability.

- **Laboratory criteria:** white blood cell (WBC) count < 4,000 or > 20,000 x 10⁹ cells/L, platelet count < 100,000 x 10⁹/L, C-reactive protein (CRP) > 10 mg/L, glucose intolerance as defined by a blood glucose value > 180 mg/dL (> 10 mmol/L) when receiving normal glucose amounts (7 – 15 g/kg/day), metabolic acidosis as defined by a base excess (BE) < –10 mmol/L (~10 mEq/L) or a blood lactate value > 2 mmol/L.
**Exclusion criteria**

Were excluded infants who had a major congenital abnormality or a chromosomal abnormality, infants with known or suspected congenital metabolic disease, neonates transfused (erythrocytes, plasma or platelets) before the collection of samples, refusal of consent.

**Samples collection**

**Blood**: 1 ml of blood was collected at birth, before any possible transfusion or therapy administration (it was picked up during blood collection for exams routinely executed at birth).

**Urine**: at least 2 ml of urine were collected from birth to 72 hours of life (urine samples were collected using a non-invasive way -a cotton swab- and subsequently transferred into a tube).

Samples were stored in a freezer at -80°C until analysis.

**Metabolomic analysis**

The analysis of samples was performed by mass spectrometry at the Laboratory of Mass Spectrometry Città della Speranza’s Institute of Pediatric Research (IRP), Women’s and Children’s Health Department of University-Azienda Ospedaliera of Padova.

**Sample analysis (untargeted metabolomics)**

Metabolic profiling was performed in positive and negative ionization mode on an Acquity UPLC system (Waters, U.K.) coupled to an QTof Synapt G2 HDMS mass spectrometer (Waters MS Technologies, Ltd., Manchester, U.K.). UPLC-MS conditions were optimized in terms of peak shape, reproducibility and retention times of analytes. Chromatography was performed using an Acquity HSS T3 (1.7 μm, 2.1 x 100 mm) and Acquity HILIC (1.7 μm, 2.1x 100 mm) column (Waters Corporation, Milford, U.S.A.) kept at 40 °C. For mass accuracy, a LockSpray interface was used with a 20 μg/L leucine enkephalin. Data were collected in continuum mode with a scan range
of 20-1200 m/z, with lockmass scans collected every 10 s and averaged over 3 scans to perform mass correction.

**Data Processing**

Data were pre-processed using Progenesis software (Waters Corporation, Milford, U.S.A.). The ion intensities for each peak detected were normalized on the basis of the regression models for the Quality Controls and, then, by Probabilistic Quotient normalization. The generated data tables, comprising m/z, RT, and intensity values for each variable in every sample, were submitted to data analysis.

**Statistical Analysis**

Both univariate and multivariate statistical data analysis were applied to characterize each group of enrolled babies. Firstly, exploratory data analysis based on Principal Component Analysis (PCA) was performed to discover outliers, trends and clusters of observations. Thus, supervised projection methods based on Projection to Latent Structures regression (PLS) were used to create mathematical models capable to discover the relationships between the measured metabolites and the outcomes recorded for the recruited infants. As a result, suitable data representations that could be interpreted in terms of the synergic effect of the measured metabolites were obtained. Specifically, post-transformation of PLS-Discriminant Analysis (ptPLS2-DA) was applied for distinguishing infants who developed sepsis from infants who did not develop sepsis on the basis of their metabolomic profiles. Robust methods for model validation were applied to avoid over-fitting and false discoveries and to guarantee the reliability of the obtained models. Univariate data analysis was used as complementary tool to integrate the results obtained by multivariate data analysis. Parametric and non-parametric tests were performed, False Discovery Rate applied to correct the significance of the tests and Receiver Operating Characteristic (ROC) curve used to describe the putative markers. Data analysis was implemented by the R 3.1.2 platform (R
Foundation for Statistical Computing) whereas PCA was performed by SIMCA 14 (MKS Data Analytics Solutions).

**Metabolites identification**

Mass spectrometry (MS) is the preeminent analytical tool for the detection and structural identification of metabolites; complex biological matrices can be analysed directly. The mass-to-charge ratio (m/z) of the ions measured by MS using LC-QTof can be readily converted into the molecular weight of each entity. Furthermore, the elemental composition of these entities can be derived if the measured m/z is sufficiently accurate (< 3 ppm). It will be the first resort to propose the metabolite identification by the consultation of the data bank of metabolite (HMDB, KEGG, PubChem, ecc.) containing confirmed and probable urine and serum/plasma metabolite.

In addition, the examination of fragmentation products ions generated from LC-MS^E or LC-MS/MS experiments can provide additional detailed structural information for the identification of the metabolites.

The validation of the metabolite identification will be made by target metabolomic analysis by triple quadrupole in LC-MS/MS experiments.

**Clinical data**

For each enrolled patient clinical data were collected as follows: demographic data (gestational age, birth weight and percentile of growth, sex, race, twins), pregnancy and delivery (antenatal steroids, twin pregnancy, hypertension, gestational diabetes, maternal fever at labor, prolonged rupture of membranes, positive vaginal swab, meconium stained amniotic fluid, obstetric complications), respiratory and cardiovascular support (ventilatory support - invasive and non-invasive ventilation, kind and duration of support-, inotropic support, pulmonary hypertension, need for surfactant therapy), sepsis (documented or suspected, clinical and laboratory presentation, microbiologic data, antibiotics or antifungal therapy), comorbidities (patent ductus arteriosus, necrotising enterocolitis,
retinopathy of prematurity, intraventricular haemorrhage, periventricular leukomalacia) and surgical interventions, length of hospital stay to first discharge home, death during or after discharge.
RESULTS

Study population

Recruitment started in March 2015 and lasted until December 2017. All subjects were recruited in the Neonatal Intensive Care Unit of Women’s and Children’s Health Department of University-Azienda Ospedaliera of Padova.

One-hundred and twenty-three subjects were enrolled in the study. Thirty neonates presented a septic episode (EOS and/or LOS): a single newborn was affected by EOS caused by E. coli (confirmed EOS), twelve neonates showed clinical and biochemical signs of EOS but their blood culture was negative (presumed EOS), thirteen babies were affected by LOS and the etiological agent was a coagulase-negative Staphylococcus (CONS) for all of them (confirmed LOS), four patients were affected by both EOS (confirmed or presumed) and LOS (confirmed or presumed). Seventeenth neonates were therefore identified as cases (confirmed or presumed EOS +/- LOS) and seventeenth gestational age-matched newborns were enrolled as controls (neither EOS nor LOS, confirmed or presumed). None of the selected neonates died. Mean gestational age was 30 weeks’ gestation (± 18.0 days SD) and 30.1 weeks’ gestation (± 16.3 days SD) for septic and non-septic infants. Mean birth weight was 1338.82 g (± 393.92 g SD) and 1297.94 g (± 333.58 g SD) for septic and non-septic neonates.

Since our Laboratory of Mass Spectrometry has a great experience in urine analysis and urine samples were more abundant than plasm, metabolomic untargeted analysis was performed on urine samples. For a few subjects, more than one sample was collected within 72 hours of birth; therefore, a total of 37 samples was analyzed. Of these, 19 (9 sepsis, 10 non-sepsis) samples were collected within 24 hours of birth, 11 (7 sepsis, 4 non-sepsis) between 24 and 48 hours after birth, 7 (4 sepsis, 3 non-sepsis) between 48 and 72 hours of age.
Urine metabolome analysis

Among the urine samples collected, those associated with sepsis (n = 20) were compared with those associated with no sepsis (n = 17). The negative data set included 2394 RT_mass variables, while the positive data set included 3224 RT_mass variables.

Univariate analysis displayed 152 variables in negative ionization mode and 600 variables in positive ionization mode with q-value < 0.10 (Fig. 1). By annotation of these signals, many putative biomarkers of neonatal sepsis could be discovered.

Fig 1. Univariate analysis results (negative data set and positive data set). Variables with q-value < 0.10 are those under the dotted red line.

Post-transformation of PLS-Discriminant Analysis (ptPLS2-DA) was applied for distinguishing infants who developed sepsis from infants who did not develop sepsis on the basis of their
metabolomic profile. Comparison of samples collected within 24 hours of birth (9 of septic neonates, 10 of non-septic neonates) showed an evident clustering, thus allowing clear differentiation between groups (Fig. 2).

Data set T3 NEG
ptPLS2-DA
(A=1+1, $R^2=0.78$, $Q^2=0.40$)

Data set T3 POS
ptPLS2-DA
(A=1+1, $R^2=0.71$, $Q^2=0.41$)

Fig 2. ptPLS2-DA model for sepsis versus non-sepsis group (negative data set and positive data set). N=non-sepsis group, Y=sepsis group.

Demographic data of the two groups (sepsis versus non-sepsis, urine collected within 24 hours of birth) were comparable (gestational age: t-test p-value=0.80, Mann-Whitney p-value=1.00; birth weight: t-test p-value=0.34, Mann-Whitney p-value=0.29; sex: Fisher exact test p-value=1.00).

Considering as metadata the gestational age, the birth weight and the sex, the PCA model on the metadata did not reveal clusters corresponding to the two groups under investigation. The PLS-DA models constructed considering the metadata as the X-block were also unreliable in modeling the differences between the two groups.
Interestingly, clustering of patients was not possible by analyzing neither samples collected between 24 and 48 hours of life nor samples collected between 48 and 72 hours of life. This suggests that the effect of antibiotic therapy, that was promptly started within the first hours of life, can nullify this differentiation.

The potential confounding influence on results of the most frequent comorbidities of premature babies (respiratory distress syndrome and patent ductus arteriosus) was evaluated and excluded (RDS: Fisher exact test p-value = 1.00; PDA: Fisher exact test p-value = 0.17).

ROC curve analysis of the 100 variables that demonstrated the best potentiality in differentiation of groups (septic versus non-septic subjects) documented high performance (AUC for positive dataset=0.88, AUC for negative dataset=0.80) (Fig. 3), comparable to those of PCR dosed between 2 and 24 hours of life (Fig. 4).

![ROC curves](image)

**Fig 3.** ROC curves (positive data set and negative data set) for metabolomics variables (t-test p-value = 0.0001-0.006; Mann-Whitney p-value = 0.0002-0.007)
Fig 4. ROC curves for PCR (2-24 hours of life)
(t-test p-value = 0.01; Mann-Whitney p-value = 0.006)

The small number of recruited subjects do not allow to evaluate the performance of a combined diagnostic tool (urine metabolome + PCR).

The metabolite allocation is in progress.
DISCUSSION

Neonatal sepsis is a complex infection-induced systemic inflammatory response syndrome and it is a main cause of mortality and morbidity in newborns. An early and accurate detection of sepsis is mandatory in neonates, since the clinical course of the infectious process can be fulminant, leading to septic shock and death within hours after the first clinical symptoms. The gold standard for diagnosis of neonatal sepsis is blood culture, but it is time-consuming and false negative results are not rare. To date, there is still no reliable biochemical marker of neonatal sepsis.

The aim of our study was to compare the metabolic profile of urine collected within 72 hours of birth between preterm neonates affected by early onset sepsis (EOS) and healthy preterm infants, searching for a specific metabolite or metabolic profile enabling the early identification of preterm newborns prone to develop EOS.

The analysis of metabolome with UPLC-MS of urine samples collected within 24 hours of birth revealed an evident clustering of subjects (septic versus non-septic neonates). This differentiation could not be highlighted in urine samples collected after 24 hours of life. This is consistent with the clinical course of disease, since the infectious process is generally well controlled if the antibiotic therapy is promptly started, and clinical improvement is generally evident within 24 hours of birth. All recruited neonates fully recovered and no one dead, so it was not possible to analyse metabolome of babies with poor prognosis. It would be interesting to compare the early (<24 hours of life) metabolic profile of poor prognosis infants with that of fully recovered neonates, to shed light to possible early metabolomic markers of prognosis. Moreover, it would be of interest to serially compare the metabolome of these groups of affected patients (poor versus good prognosis) and healthy newborns, to evaluate if metabolic excursions from normal equally recover in both the affected infants’ groups or, as expected according to clinical course, it does not.

Our results show a diagnostic performance of urine metabolome comparable to those of PCR, as revealed by the ROC curves. The small number of recruited subjects did not allow to evaluate the performance of a combined diagnostic tool (urine metabolome + PCR), but it presumably would be
excellent and this possibility will be investigated in future researches. The annotation of the variables that demonstrated the best potentiality in differentiation of groups (septic versus non-septic subjects) will allow to identify many putative biomarkers of neonatal sepsis. Some of these metabolites could show optimal characteristics to be qualified as an ideal biomarker (high sensitivity and specificity, small blood volume, easy laboratory detection, reasonable price). Most of all, the identification of these variables will shed light to the complex biologic pathways that underlie neonatal sepsis, allowing to understand its mechanisms of evolution. Some of the identified metabolites could appear as a good tool in monitoring response to antibiotic therapy or to predict the evolution of the disease.

To date, only two studies analyzing metabolomics perturbations in neonatal sepsis were published. The study by Fanos et al. reported the presence of a specific urine metabolic profile in 9 septic neonates (both EOS and LOS) when compared to 16 non-affected infants. The molecules responsible for the differences in the metabolic profiles were glucose, lactate and acetate, the urine content of which was increased in septic neonates compared to controls, while 2,3,4-trihydroxybutyric acid (THBA), ribitol, ribonic acid and citrate concentration was decreased. Despite the small number of recruited subjects and the difference in gestational age between groups (29.1 weeks’ gestation for septic neonates and 34.6 weeks’ gestation for non-affected newborns), this study firstly revealed the difference between metabolic profile of septic neonates and controls, with the purpose of defining the metabolic patterns associated with this pathology.71

Sarafidis et al. revealed that the metabolic profile of neonates with proven and possible late onset sepsis (LOS) was markedly different compared to those of infants without sepsis. The metabolic alterations seen were mainly relevant to energy producing biosynthetic pathways and basic structural components of the organism. This was the first metabolomic study where subjects were serially evaluated at three time-points during the acute phase and upon recovery from the disease. This allowed to determine that less discrimination and no discrimination between metabolome of
septic and non-septic neonates could be made respectively on the day 3 and 10 of sepsis, suggesting that metabolomics could be a reliable tool to monitor response to therapy.\textsuperscript{72}

As far as we know, our research is the largest metabolomic study performed in neonates affected by early onset sepsis. EOS is a major concern for both term and preterm neonates. Its etiology reflects transplacental or ascending infections from the maternal genital tract and symptoms are frequently severe with fulminant course, differently from most of LOS, unless a broad spectrum antibiotic therapy is promptly started. This explains our primary clinical interest in investigating EOS.

Moreover, since our final aim is the search for an accurate biomarker of neonatal sepsis by untargeted metabolomic approach and this methodology is easier if no confounding factors deriving from concomitant therapies or comorbidities are present, in our opinion EOS is a field of investigation more suitable than LOS for this purpose. If annotation and identification of variables discriminated by untargeted metabolomic analysis will allow to find a single metabolite or profile of metabolites that could be used as sepsis biomarker, its usefulness will be applied to LOS too.

Revealing a clear clustering of metabolome of septic neonates compared to healthy newborns, our results confirm published data.\textsuperscript{71,72} Interestingly, our analysis supports also the precedent finding of Sarafidis, who reported that the metabolic discrimination between septic and non-septic infants decreases until it disappears along with the clinical and laboratory improvement of sepsis.\textsuperscript{72} These results shed light to the potential of metabolomics to elucidate the biological pathways and pathophysiological mechanisms of sepsis, offering opportunities for diagnostic and therapeutic intervention.

Another strength of our work is the low gestational age of recruited infants (mean GA: 30 weeks’ gestation), significantly lower than those of neonates enrolled by Sarafidis et al. (34.7 ± 3.6 weeks for septic neonates and 35 ± 3.2 weeks for controls). Clinical presentation is a milestone of sepsis diagnosis in newborns, but signs and symptoms are not-specific and this is even more worrying at low gestational ages, when comorbidities such as respiratory distress syndrome (RDS) and patent ductus arteriosus (PDA) may mimic sepsis. Therefore, primarily in the most premature babies the
discovery of new biochemical markers that could aid in formulating the diagnosis of sepsis is warranted.

Our study is the first to consider the potential confounding influence on results of the most frequent comorbidities of premature babies (RDS and PDA), excluding a role of these diseases. This aspect is of great importance, since previously untargeted metabolic profiling of unrelated diseases yielded similar findings among patients with different diseases as compared to healthy controls.\textsuperscript{75}

A limit of our study is the lack of variables annotation and metabolite identification. Nevertheless, these further investigations are ongoing. The ultra-sensible technique used to perform the analysis is a great strength of this research. It allowed to dose a great number of variables (>5000) and potential putative biomarkers (>750). We expect to identify metabolites involved in biological pathways, even unexplored, strictly related to sepsis evolution. This will increase the chances of finding a reliable and specific biomarker of sepsis and of exploring the pathophysiology of such a complex disease.

When the identification of putative biomarkers (single metabolites or metabolites profiles) will be completed, a targeted metabolomic analysis will be performed on plasm samples collected from the same recruited neonates, to investigate their presence in this different biologic fluid.

Finally, the recruitment of a validation population is ongoing, to confirm the results.

Future progresses of this research could be the application of the results to term neonates and late onset sepsis events. Particularly interesting would be also the identification of biomarkers associated with a worse course of the disease, thus enabling the development of a prognostic score for neonatal sepsis.
CONCLUSIONS

In conclusion, neonates with early onset sepsis showed a specific metabolic profile compared to those of newborns not affected by sepsis at the onset of infection, allowing their clear discrimination with the use of UPLC-MS-based urine analysis.

The diagnostic performance of urine metabolome resulted comparable to those of PCR.

Results of this research support the effectiveness of metabolomics in exploring biochemical pathways of neonatal sepsis, potentially providing novel putative biomarkers for early diagnosis, guide to antibiotic therapy, and monitoring of disease progression.

Future perspectives will be addressed to introduce these novel metabolomics biomarkers in clinical routine practice, ensuring short laboratory turnaround time and 24 hours bedside availability.
REFERENCES

1. WHO 2016, www.who.it/gho/data


