UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede amministrativa: Università degli Studi di Padova

Dipartimento di Salute della Donna e del Bambino

CORSO DI DOTTORATO DI RICERCA IN:
Medicina dello Sviluppo e Scienze della Programmazione Sanitaria

Curriculum: Emato-Oncologia, Genetica, Malattie Rare e Medicina predittiva

Ciclo: XXXIV°

Untargeted and targeted metabolomic approach for the study of neonatal sepsis

Approccio metabolomico untargeted e targeted per lo studio della sepsi neonatale

Coordinatore: Ch.mo Prof. Gianni Bisogno
Supervisore: Ch.mo Prof. Eugenio Baraldi
Tutor: Dott. Giuseppe Giordano

Dottorando: Gabriele Poloniato

Summary

1 Abstract

1.1 English Abstract

Sepsis is a major concern in neonatology. Neonatal sepsis is an infection-induced, systemic inflammatory response syndrome common in premature and term neonates. It is one of the leading causes of neonatal death and morbidity and is believed to have a key role in most inflammatory disorders that cause or enhance the main morbidities affecting the preterm (bronchopulmonary dysplasia, white matter injury, necrotizing enterocolitis, and retinopathy of prematurity). Sepsis in the newborn is typically classified as either early-onset sepsis (EOS), when the infection occurs within three days after birth, or late-onset sepsis (LOS) if it develops afterward. Early detection of neonatal sepsis and prompt administration of broad-spectrum antibiotic therapy can prevent its clinical course towards septic shock and death, but it is not easy to diagnose neonatal sepsis early on. Blood culture is still considered the gold standard, even though it takes time to obtain the results, and false-negative findings are not uncommon because neonatal bacteremia is often intermittent, and intrapartum antibiotic treatment may limit the culture's diagnostic value. Neonatal sepsis is therefore mainly suspected on the grounds of non-specific clinical signs and symptoms; moreover, none of the most widely used biomarkers are entirely reliable indicators of sepsis in newborns. Hence, identifying new biomarkers for EOS is of crucial importance.

Furthermore, while supportive therapies promote the survival of septic neonates, there are no mechanistic therapies to alter the underlying pathophysiology, and this is partly due to partial knowledge of the complex biological pathways underlying the pathophysiology of sepsis.

The aim of the study was to compare the metabolic profiles of plasma and urine samples collected at birth from preterm neonates with and without early-onset sepsis (EOS) to identify metabolic perturbations that might orient the search for new early biomarkers.

All preterm newborns admitted to the neonatal intensive care unit were eligible for this proof-of-concept, prospective case-control study. Infants were enrolled as "cases" if they

developed EOS, and as “controls” if they did not. Plasma samples collected at birth and urine samples collected within 24 h of birth underwent untargeted and targeted metabolomic analysis using mass spectrometry coupled with ultra-performance liquid chromatography. Univariate and multivariate statistical analyses were applied. Of 123 eligible newborns, 15 developed EOS. These 15 newborns matched controls for gestational age and weight.

UPLC–MS analysis of urine samples revealed a clustering of cases of EOS compared with healthy neonates. Furthermore, a metabolic signature exists to distinguish neonates that develop sepsis and healthy subjects and putative markers discriminating between EOS cases and controls were discovered. Pathway analysis showed metabolic derangements most involved in EOS. The most significant metabolic pathways were investigated using a targeted analysis on plasma samples collected from the same neonates, confirming the marked disruption of the tryptophan and glutathione metabolic pathways in the neonates with EOS.

In conclusion, neonates with EOS had a metabolic profile at birth that clearly distinguished them from those without sepsis, and metabolites of glutathione and tryptophan pathways are promising as new biomarkers of neonatal sepsis.

1.2 Italian Abstract

La sepsi è una delle principali problematiche in ambito neonatologico. La sepsi neonatale è una sindrome da risposta infiammatoria sistemica indotta da infezione comune nei neonati prematuri e a termine. È una delle principali cause di morte e morbilità neonatale e si ritiene che abbia un ruolo chiave nella maggior parte dei disturbi infiammatori che causano o contribuiscono allo sviluppo delle principali morbilità che colpiscono il prematuro (displasia broncopolmonare, danno della sostanza bianca cerebrale, enterocolite necrotizzante e retinopatia del prematuro). La sepsi nel neonato è solitamente classificata come sepsi ad esordio precoce (EOS), quando l'infezione si manifesta entro tre giorni dalla nascita, o sepsi ad esordio tardivo (LOS) se si sviluppa in seguito. La diagnosi precoce della sepsi neonatale e la pronta somministrazione di una terapia antibiotica ad ampio spettro possono prevenirne il decorso clinico verso lo shock settico e la morte, ma non è facile diagnosticare precocemente la sepsi neonatale. L'emocoltura è ancora considerata il gold standard, anche se ci vuole tempo per ottenere i risultati e i falsi negativi non sono rari perché la batteriemia neonatale è spesso intermittente e il trattamento antibiotico intrapartum può limitare il valore diagnostico della coltura. La sepsi neonatale viene quindi ipotizzata principalmente sulla base di segni e sintomi clinici non specifici; per di più, nessuno dei biomarcatori più utilizzati è un indicatore del tutto affidabile di sepsi nei neonati. Pertanto, l'identificazione di nuovi biomarcatori per EOS è di cruciale importanza. Inoltre, mentre le terapie di supporto promuovono la sopravvivenza dei neonati settici, non esistono terapie per alterare la fisiopatologia sottostante, e ciò è in parte dovuto alla parziale conoscenza delle complesse vie biologiche alla base della fisiopatologia della sepsi.

Lo scopo dello studio era confrontare i profili metabolici di campioni di plasma e urina raccolti alla nascita da neonati pretermine con e senza sepsi ad esordio precoce (EOS) per identificare le perturbazioni metaboliche che potrebbero orientare la ricerca di nuovi biomarcatori precoci.

Tutti i neonati prematuri ammessi all'unità di terapia intensiva neonatale erano eleggibili per questo studio prospettico caso-controllo. I neonati sono stati arruolati come "casi" se hanno sviluppato EOS e come "controlli" in caso contrario. I campioni di plasma raccolti alla nascita e i campioni di urina raccolti entro 24 ore dalla nascita sono stati sottoposti ad

analisi metabolomica targeted e untargeted tramite spettrometria di massa accoppiata a cromatografia liquida UPLC. Sono state poi applicate analisi statistiche univariate e multivariate. Di 123 neonati idonei, 15 hanno sviluppato EOS. Questi 15 neonati corrispondevano ai controlli per età gestazionale e peso.

L'analisi UPLC-MS dei campioni di urina ha rivelato un cluster di casi di EOS rispetto ai neonati sani. Inoltre, esiste una differenza metabolica per distinguere i neonati che sviluppano sepsi dai soggetti sani e sono stati scoperti marcatori putativi che discriminano tra casi di EOS e controlli. L'analisi dei pathways ha evidenziato gli squilibri metabolici maggiormente coinvolti in caso di EOS. Le vie metaboliche più significative sono state studiate utilizzando un'analisi targetd su campioni di plasma prelevati dagli stessi neonati, confermando l'evidente perturbazione delle vie metaboliche del triptofano e del glutatione nei neonati con EOS.

In conclusione, i neonati con EOS presentavano un profilo metabolico alla nascita che li distingueva chiaramente da quelli senza sepsi, e i metaboliti delle vie del glutatione e del triptofano sono promettenti come nuovi biomarcatori della sepsi neonatale.

2 List of abbreviations

Abbreviation	Meaning
1D NMR	1 dimensional proton nuclear magnetic resonance
^{1}H-NMR	Proton nuclear magnetic resonance
2D NMR	2 dimensional proton nuclear magnetic resonance
ANC	Absolute neutrophil counts
ANOVA	Analysis of Variance
APCI	Atmospheric pressure chemical ionization
APPI	Atmospheric pressure photo-ionization
AUC	Area under the curve
BW	Birth weight
CAP	Community-acquired pneumonia
CCS	Collision cross section
CE	Capillary electrophoresis
CI	Confidence interval
CONS	Coagulase-negative staphylococci
CRP	C-reactive protein
CSF	Cerebrospinal fluid
CTRL	Control
DI-MS	Direct infusion-mass spectrometry
DOSY	Diffusion ordered spectroscopy
DRMPs	Death-related metabolic pathways
EI-MS	Electron impact ionization-mass spectrometry
EOS	Early-onset sepsis
ESI	Electrospray ionization
GA	Gestational age
GBS	Group B *Streptococcus*
GC	Gas chromatography
GC-EI-MS	Gas chromatography-electron impact ionization-mass spectrometry
GC-MS	Gas chromatography-mass spectrometry
HILIC	Hydrophilic Interaction Liquid Chromatography
HMBC	Heteronuclear multiple bond correlation
HMDB	Human Metabolome Database
HPLC	High performance liquid chromatography
HR-MS-MS	Tandem high resolution mass spectrometry

$HRMS^n$	Multistage high resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
I/T	Immature to total neutrophils ratio
IAP	Intrapartum antibiotic prophylaxis
ICR	Ion cyclotron resonance
ICU	Intensive care unit
Il-1	Interleukin-1
Il-6	Interleukin-6
Il-8	Interleukin-8
IM	Ion mobility
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-ESI-MS	Liquid chromatography-electrospray ionization-mass spectrometry
LC-HRMS	Liquid chromatography-high resolution mass spectrometry
LC-IM-MS	Liquid chromatography-ion mobility-mass spectrometry
LC-MS	Liquid chromatography-mass spectrometry
LIT	Linear quadrupole ion
LLE	Liquid phase extractions
LOI	Late onset infection
LOQ	Limit of quantification
LOS	Late-onset sepsis
m/z	Mass to charge ratio
MCC	Matthew correlation coefficient
MCC5-fold	Matthew correlation coefficient calculated by five-fold cross-validation
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS-MS	Tandem mass spectrometry
MS^n	Multistage mass spectrometry
MW	Molecular weight
NEC	Necrotizing enterocolitis
NEG	Negative
NICU	Neonatal intensive care unit
NMR	Nuclear magnetic resonance
nSOFA	Neonatal sequential organ failure assessment
OPLS-DA	Orthogonalartial least squares discriminant analysis
PC	Principal component
PCA	Principal components analysis
PCR	Polymerase chain reaction
PCT	Procalcitonin

PICU	Pediatric intensive care unit
PLS	Partial least squares
PLS-DA	Partial least squares discriminant analysis
POS	Positive
PTNB	Preterm newborns
QC	Quality control
qPCR	Quantitative polymerase chain reaction
QTOF	Quadrupole time-of-flight
RT	Retention time
SAA	Serum amyloid A
SIRS	Systemic inflammatory response syndrome
SOFA	Sequential organ failure assessment
SPE	Solid phase extractions
SSN	Septic shock non-survivor
SSS	Septic shock survivor
TLC	Thin layer chromatography
TNFa	Tumor necrosis factor-alpha
TOCSY	Total correlation spectroscopy
TOF	Time-of- flight
UPLC	Ultra performance liquid chromatography
UPLC-MS	Ultra performance liquid chromatography-mass spectrometry
VIP	Variable influence on projection
VLBW	Very low birth weight
WBC	White blood cell

3. Background

3.1 Sepsis

A first definition of sepsis, derived from a Greek term used by Hippocrates to describe natural decay of organic matter, was given in 1914 by Hugo Schottmüller who wrote that "sepsis is present if a focus has developed from which pathogenic bacteria, constantly or periodically, invade the blood stream in such a way that this causes subjective and objective symptoms" [1, 2]. Nowadays, sepsis is defined by Sepsis-3 conference as a "life-threatening organ dysfunction caused by a deregulated host response to infection", so consisting in a dysfunction affecting different organs caused by infectious microorganisms and induced by mediators of inflammation and dysregulated host response to infection, that cause alteration of the immune, inflammatory and coagulative equilibrium and its more severe forms, such as septic shock, may lead to organ dysfunction, organ failure, and death [3, 4, 5, 6].

Sepsis is a main cause of mortality and morbidity in newborns, with an estimated global incidence of 3 million cases in neonates, and it can be classified as early onset sepsis (EOS), which takes place in the first 72 hours and late onset sepsis (LOS), which begins after 72 hours and is usually caused by nosocomial pathogens [4, 7, 8].

Early onset sepsis is a major cause of mortality and morbidity in neonates, particularly preterm newborns and those with a very low birth weight (VLBW) and, although its incidence is lower than in the past, EOS is still a problem [9, 10]. It is acquired in utero through intraamniotic infection, rupture of membranes, or during passage through the vagina, with *Escherichia coli, Streptococcus agalactiae* and Group B *Streptococcus* (GBS) in general as major responsible pathogens [9, 11, 12]. In a study, Weston et al. reported that the most commonly pathogens were GBS (37.8%), *E. coli* (24.2%), viridans Streptococci (17.9%), *Stapylococcus aureus* (4.0%), *and Haemophilus influenzae* (4.0%) [13]. Similarly, Stoll and colleagues outpointed GBS (43%) as the most frequent early onset pathogens, followed by *E. coli* (29%), but infection with *S. aureus* was infrequent, and that rates of *E. coli* infection were higher than rates of GBS among very low birth weight infants

(5.09 vs 2.08 per 1000 LBs) and infants with a BW of 1501 to 2500 g (0.54 vs 0.38 per 1000 LBs) while, in infants with a BW of >2500 g, GBS rates were higher than *E. coli* rates (0.35 vs 0.07 per 1000 LBs) [14].

Nowadays, the incidence of EOS in infants born at term is of approximately 0.5/1000 up to 0.8/1000 live births, but this number doubles in late preterm newborns (PTNB), and even more significant in PTNB<34 weeks and NB with VLBW registering values of 15 to 19 per 1000 live births; however, the incidence of EOS varies by country and also between different regions [15, 16, 17].

GBS intrapartum prophylaxis and implementation of universal GBS screening in 2012 has resulted in a notable reduction in the invasive EOS burden and identify interventions to specifically prevent early onset *E. coli* infections is a promising way to reduce EOS burden; furthermore, identification and implementation of strategies to prevent preterm birth, such as prenatal care and risk directed interventions, would offer further hope for substantial reductions in EOS [11, 12, 13, 15].

Apart from maternal GBS colonization, demographic factors, preterm birth, rupture of membrane for more than 18 hours and maternal signs of intraamniotic infection are all risk factors for EOS [18].

From 2017 to 2019, the American College of Obstetricians and Gynecologists and the American Academy of Pediatrics updated guidance for intrapartum antibiotic use in women with concern for evolving intraamniotic infection, for antenatal screening and intrapartum antibiotic prophylaxis (IAP) to prevent GBS infection and for administration of empirical antibiotic therapy to newborns at risk for EOS. However, innovative clinical and public health approaches to prevent EOS are urgently needed, including efforts to prevent maternal intraamniotic infection [19].

LOS onset is most frequently defined at 72 h after birth, a time that allows to differentiate LOS from EOS in terms of the spectrum of causative pathogens. LOS is more common among the most premature infants, and it's associated with the postnatal nosocomial or community environment; in fact, incidence of LOS has increased in parallel with the improved survival of premature infants, especially in VLBW ones, indicating the role of hospitalization and life sustaining medical devices in the pathogenesis [20, 21, 22, 23, 24].

Incidence of LOS in the neonatal intensive care unit (NICU) varies according to birth weight (BW) and gestational age (GA) and ranged from 25–30% in very low birth weight

(VLBW: ≤1500 g) infants1,3 to 6.2–10% in late-preterm (GA: 34–37 weeks) infants [25]. Apart from immaturity, other risk factors for LOS include patent ductus arteriosus, necrotizing enterocolitis (NEC), the long-term use of invasive interventions, such as mechanical ventilation and intravascular catheterization, the failure of early enteral feeding with breast milk, a prolonged duration of parenteral nutrition, hospitalization, surgery and underlying respiratory and cardiovascular diseases, but also delivery outside the health facility [20, 25, 26, 27].

Gram-positive organisms are the most common pathogens causes of LOS, where coagulase-negative staphylococci (CONS) have emerged as the predominant pathogens of LOS, accounting for 53.2%–77.9% of LOS in industrialized countries and 35.5%–47.4% in some developing regions; however, CONS are not as virulent as Gram-negative bacteria and possibly fungi. Other Gram-positive organisms responsible of LOS are *Enterococcus* spp., Group B *Streptococcus, Streptococcus pneumoniae* and *Viridans Streptococcus* [20, 25, 28, 29]. Gram-negative organisms, including *Klebsiella pneumoniae, Escherichia coli, Enterobacter aerogenes* and *Pseudomonas aeruginosa,* are usually the second cause of LOS, followed by fungus (*Candida albicans, Candida parapsilosis*) and polymicrobial microorganisms [20, 25, 30, 31].

Given the morbidity and mortality associated with LOS, several strategies have been implemented in NICUs worldwide to decrease the incidence of infection, such as maternal chemoprophylaxis for prevention of early-onset group B streptococcal infection, rigorous hand hygiene procedures, improved central line care with central line bundles, antifungal prophylaxis, and careful attention to NICU design and staffing [29]. However, application of broad-spectrum antibiotics in the past decades has contributed to an increasing incidence of multidrug-resistant Gram-negative bacilli [20] and bacteria in general.

Rapid identification of newborns with infection is a daily challenge for pediatricians worldwide. Symptoms of EOS such as respiratory distress, hypo- or hyperthermia or feeding intolerance are non-specific as these symptoms are also often observed in neonates without infection. Unnecessary treatment and hospitalization is undesirable because of the risk of promoting multidrug resistant bacteria. Furthermore, antibiotic treatment early in life disturbs the microbial flora colonizing the neonate and is associated with important health problems such as eczema, allergies, and inflammatory bowel diseases in later life. Moreover, financial costs and use of resources due to unnecessary hospitalization are to be

considered [32, 33].

The challenge of fast diagnosis of sepsis is that this syndrome is based on highly complex pathophysiological pathways that may show varying clinical signs and symptoms, common to other clinical conditions [34, 35].

Due to the many non-specific signs of sepsis, laboratory tests are helpful to diagnose sepsis and distinguishing it from other conditions, by isolating causative agent from a normally sterile blood site (blood, urine, CSF) [6, 34, 36].

Blood culture is considered the definitive diagnostic tool for neonatal sepsis, with preferably a minimum of 0,5-1 mL of blood from two different venipunctures from two different sites, but it is time consuming requiring long time for the results and false positive results due to contaminations are possible. All blood cultures are incubated and monitored for 72 h in order to ascertain the condition of sepsis into the infants. Modified and advanced blood culture systems, like, BACTEC and BACT/ALERT are also available, which allows early and rapid screening of these blood cultures. Moreover, these cultures usually have low sensitivity towards identification of the neonatal sepsis, since small inoculation volumes in culture bottles, the use of intrapartum antibiotics and low degree of neonatal bacteraemia may cause the decreased sensitivity [20, 34, 37, 38].

Recently, molecular-based methods have emerged as promising diagnostic tools for neonatal sepsis. PCR, a technology based on the extraction of microbial DNA from blood samples and the subsequent sequencing or hybridization of species-specific gene regions, is widely investigated for the detection of micro-organisms. Furthermore, real-time PCR has been explored to monitor the microbial load and rapidly target specific micro-organisms in clinical specimens. Compared with the conventional culture technology, PCR presents a higher sensitivity, needs smaller sample volume and less laboratory time and identifies almost twice the number of positive specimens compared with conventional blood culture [39]. Recently developed PCR-based diagnostic platforms are characterized by a low contamination rate, with DNA extraction, multiplex PCR and detection of PCR products performed in a closed system, helping to differentiate potential contamination from true positive cases, particularly for the detection of CONS [20, 34, 39]. The detection of microbial pathogens in blood or cerebrospinal fluid by molecular assays based on PCR seems a promising replacement for conventional culturing. Interestingly, in contrast with what reported before, a recent systematic review and meta-analysis showed that these

assays lack sufficient sensitivity to replace microbial cultures to diagnose EOS [32]. Because PCR is a highly sensitive and rapid technique, it is increasingly being applied to bodily fluids directly without the need to first culture causative agents. Quantitative real-time amplification systems (qPCR) have a very high negative predictive value and results are usually available in a timely manner, but presents the inability to do susceptibility testing and a high sensitivity that does not differentiate between active infection and recent infections already resolved. The possibility of detecting contaminants is also high, and therefore clinical correlation with results is mandatory [34].

However, nowadays neonatal sepsis is mainly suspected based on clinical observation, evaluating several symptoms including fever, hypotension, metabolic acidosis, tachycardia or bradycardia, apnoea, respiratory distress, irritability, lethargy, feeding intolerance, abdominal distension and bleeding, that can be common also in neonates without infection and, furthermore, the disease may be present without the appearance of clinical symptoms [32, 34, 35, 38].

Sepsis-3 definitions replace cryteria for systemic inflammatory response syndrome (SIRS) in identifying patients with sepsis, but they are non-specific and not always present in patients with infection. Part of the new Sepsis-3 definitions is SOFA [Sequential (Sepsis-Related) Organ Failure Assessment, or SOFA score] as a grading score for defining acute organ dysfunction, that allocates points according to pathological change in six different organ systems: an increase in the total SOFA score by at least two points indicates acute organ dysfunction, and the diagnosis of sepsis is met if an infection is identified in parallel [6, 35, 40]. In adults, SOFA operationalizes mortality risk with infection and defines sepsis, and similarly neonatal SOFA (nSOFA) has been developed to establish a definition of sepsis for newborns. Among preterm VLBW neonates confirmed with late onset infection LOI, the nSOFA showed generalizable utility as an operational definition of organ dysfunction associated with mortality risk and may provide the requisite foundation on which to build a consensus definition for sepsis in preterm neonates [41, 42].

Given the insensitivity of physical examination and culture, neonatal sepsis is usually diagnosed on a combination of clinical signs in association with laboratory markers that include blood counts and acute-phase reactants and major efforts have been made to find biomarkers that allow early diagnosis of this disease. [4, 6, 43].

C-reactive protein CRP is an acute-phase protein synthesized by hepatocytes when the body

is affected by microbial invasion or tissue damage and the most studied infection and inflammation marker. CRP concentration in healthy patients remains stable, indicating the absence of sepsis, but during infection by pathogens, its level increases within 6-10 h and peaks after 36-50 hours. The value of CRP for the diagnosis of sepsis patients is a moderate degree: in newborn infants of <28 days of age with suspected sepsis, CRP was independently predictive of a positive blood culture and in 176 newborns >1,500 g, CRP was calculated to have a negative predictive value of 99%. [6, 38, 43, 44, 45]. Moreover, the CRP level within 48 h before treatment has been shown to potentially help in assessing the response of patients with sepsis to initiate antimicrobial therapy, and CRP level at admission could be used as a possible marker of early infection. During treatment, CRP levels begin to decline, and this might be as a result of the remission of inflammation and as well as response to antibiotic treatment, but continues to contribute to the prognosis and treatment progress monitoring in the case of sepsis, and elevated CRP levels might have a relation to the extent of the disease and the severity of the infection [44].

Furthermore, several antenatal and perinatal variables like duration of active labor, prenatal steroids, time of ruptured membranes, and intrapartum antimicrobial prophylaxis had a significant effect on CRP [46]. CRP is quite unspecific and it does not differentiate sepsis from other diseases, but it is commonly used to screen for early onset neonatal sepsis because its sensitivity has been shown to be very high, resulting useful in decreased the duration of antibiotic therapy and hospital stay, and hence reduced healthcare costs [6, 47].

Procalcitonin (PCT), a precursor of the hormone calcitonin secreted by C cells from the thyroid gland, is more specific than CRP for bacterial infections and rises more rapidly in response to infection than CRP [44, 48]. During sepsis, PCT expression increased significantly within 2 to 6 h and peaked at 6 to 24 h [44].

Also, PCT in cord blood had shown high likelihood ratio (5.72) and sufficient sensitivity (82-92%) and specificity (86-97%) to be considered a reliable rule-in and rule-out test. Thus, it has been shown the usefulness of procalcitonin-guided decision making, even superior to standard care, in reducing antibiotic therapy in neonates with suspected EOS [38, 43, 48, 49]. Thus, based on the comprehensive analysis of all biological markers of inflammatory response, PCT is considered to be the best choice among the recommendations of the guidelines and thus, PCT is used as one of the serum markers for the early diagnosis of sepsis or septic shock, and displays a higher sensitivity and specificity

than traditional serum markers [44].

PCT levels, however, can be elevated with non-infectious conditions such as respiratory distress syndrome, pneumothorax, intracranial hemorrhage, and hemodynamic instability, birth asphyxia and neonatal hypoxemia [38, 48]. Moreover, it seems that PCT rise in healthy preterm neonates happens earlier than in healthy term newborns and it also seems to be higher and longer too [46]. These aspects could weaken PCT specificity.

CD64 (FcγR1) is a high affinity receptor that belongs to a family of immunoglobulins expressed mainly on macrophages and monocytes, and is also a key immunomodulator in innate and adaptive immune responses. In healthy volunteers, FcγR1/CD64 is expressed at extremely low levels on neutrophils, but is significantly elevated after inflammation or infection [44, 50]. FcγR1/CD64 expression on neutrophils is also reported to be associated with the severity of SIRS and sepsis, resulting was a useful biomarker for the early diagnosis of sepsis adult patients [44].

The role of CD64 in neonatal sepsis has been investigated in the past with promising results. Increased expression of CD64 can be detected within 1-6 hours of bacterial invasion, and the levels remain elevated for up to 24 hours and it has shown higher concentrations in septic newborns than healthy subjects with a high specificity to diagnose LOS in preterm newborns. Thus, CD64 has limitation as single marker, but it might be useful if used in combination with CRP, guiding the decision to continue antibiotic therapy after 36-48 hours [50, 51, 52, 53].

Serum Amyloid A (SAA) is an Apolipoprotein, which is synthesized by liver and it is an acute phase reactant with a hepatic synthesis, regulated by proinflammatory cytokines, used as markers for acute and chronic inflammation. SAA levels are controlled by interleukin-1, interleukin-6 and tumor necrosis factor-alpha TNFa [38, 50, 54]. SAA is secreted in response to injury or infection; in neonatal sepsis patients serum levels of SAA have registered a significantly increase compared to healthy patients, even 1000 times higher, and when compared to CRP, SAA levels rise more rapidly, peak on day 3 of sepsis, and return to baseline levels after four days. SAA exhibits specificity of 95% and sensitivity of 82%, both higher than CRP. SAA levels is dependent on host nutritional status and hepatic function. SAA levels have higher accuracy in the early detection of neonatal infections and are inversely related to mortality in neonates [38, 50, 54, 55].

Therefore, SAA protein could help the clinicians to diagnose most cases of neonatal sepsis

and kinetically, it looks promising enough to be used as a solitary marker, but further studies need to be performed. Furthermore, gestational age and birth weight effects on SAA secretions in response to sepsis still need to be evaluated and optimal cut-off values as neonatal sepsis marker would need to be established [55, 56].

Presepsin is a soluble CD14 subtype that is released from the surface of immune cells due to stimulation by pathogens. Soluble CD14 is secreted by hepatocytes, suggesting it may be useful as an acute-phase reactant [50]. Among the new emerging biomarkers of sepsis, presepsin appears to be the most promising. Several studies have shown that presepsin plasma levels increase during bacterial sepsis and decline in response to appropriate therapy, with sensitivity and specificity values comparable, or even superior, to those of PCT and CRP with both high sensitivity of 91-94% and high specificity of 91-100%. Notably, this marker has been shown to be a reliable diagnostic and prognostic marker of adult sepsis. [50, 51, 57, 58]. In neonatal sepsis, presepsin compared to PCT has been shown to be more effective in diagnosing and guiding therapy. Since in sepsis its plasma levels increase before those of PCT and since the current methods available allow measurement of presepsin plasma levels within 17 min, it appears a sepsis biomarker particularly suited to the emergency department and critical care [50, 57].

However, presepsin measurement is not ubiquitously available as a routine laboratory test unlike other biomarkers of sepsis and it does not appear to be clearly superior to the biomarkers commonly used in the assessment of sepsis, but may be valuable when used in conjunction with other established tests to better identify patients at risk of clinical deterioration. In addition, identification of reliable cutoff levels and testing of its utility in clinical decision-making is required before use becomes more widespread. Further studies are indicated to establish whether it is useful for predicting the most severe complications of sepsis in the intensive care setting [51, 58].

Absolute neutrophil counts (ANC), white blood cell (WBC) counts and the ratio of immature to total neutrophils (I/T) in the blood are the other diagnostic non-specific tests for the identification and diagnosis of neonatal sepsis. WBC counts give a poor positive indicative value for sepsis, while neutrophil values depend upon the individual's age and neutropenia can serve as a sensitive measure but depends upon delivery method, gestational age and altitude [38]. WBC and ANC demonstrated to be most informative when low (WBC <5,000 and ANC <1,000 at ≥ 4 h had likelihood ratios of 81 and 115, respectively)

and no test was very sensitive. Anyway, neutrophil counts can also be affected by maternal (hypertension and fever), intrapartum (asphyxia, meconium aspiration syndrome, type of delivery), and perinatal (periventricular hemorrhage, reticulocytosis, hemolytic disease and pneumothorax) factors. Diagnostic accuracy of WBC indices improved with advancing postnatal age, and was most useful after 4 hours of life [43, 51]. Additionally, Platelet counts are also carried out as preliminary diagnostic testing for neonatal sepsis, but it's not very specific towards neonatal sepsis [38].

I/T ratio is a promising useful tool for early onset sepsis (EOS) with reasonable specificity but cannot be relied upon as sole indicator. Combination of normal immature to total neutrophil Ratio with negative CRP values in neonates with presumed sepsis demonstrated to be an indicator of non-infected neonate [59].

The combination of multiple biomarkers, such as the total number of neutrophils, immature to total neutrophil ratio and C-reactive protein (CRP) and clinical findings, holds promise to increase the sensitivity and to enable a fast and accurate diagnosis of neonatal sepsis [20, 43, 51] and serial measurements of CRP combined with other acute phase reactants such as procalcitonin, IL-6, and IL-8 may improve its diagnostic accuracy [48].

However, there is a lack of highly sensitive and specific diagnostic tools: conventional laboratory parameters as C reactive protein (CRP), and white blood cell count (WBC) are non-specific, and not adequately sensitive [32].

So, early diagnosis of neonatal sepsis remains difficult and several studies have been conducted in order to identify a test satisfying criteria to make it the ideal marker for sepsis in neonates [4].

Several other biomarkers (endocan, cytokines, chemokines, lipopolysaccharide binding protein, hepcidin, sTREM1...) have been recently studied for sepsis, but none of them is considered valid for the diagnosis. Thus, there is still no single biomarker identified for neonatal sepsis diagnosis, and consensus on an ideal biomarker will be vital to aid proper diagnosis and adequate treatment, which should reduce mortality levels associated with neonatal LOS [50, 51, 54].

Furthermore, as previously discussed, there is still no currently available test that is able to perfectly diagnostic accuracy. However, combinations of biomarkers with diagnostic tests could enable to obtain an accurate diagnosis of neonatal sepsis.

3.2 Metabolomics

In the last decades, the advances in Nuclear Magnetic Resonance (NMR), Mass Spectrometry (MS) and chromatographic instruments have contributed to the development of systems biology, that focuses on complex interactions within biological structures using a holistic perspective approach [60]. One of the principal scientific progresses in this field was the development of the so called "-omics" sciences. These are technologies that allow to characterize several classes of biological molecules, such as genes, proteins and metabolites, obtaining information about the functional activity of biochemical pathways, and of the structural genetic differences [61, 62]. The sequencing of DNA maps gave birth to genomics, the first of these "-omics" techniques, concerning the characterization of genetic material [63]. Transcriptomics is defined as the study of gene expression, describing the full set of mRNA present in a cell or tissue (transcriptome) and, at the same way, proteomics studies the protein translation, and the proteome is the complete set of proteins expressed in a cell or tissue at a time [64, 65].

The term "metabolomics" was used for the first time in a scientific paper in 1998 by Steven Oliver [66], and it could easily be confused with "metabonomics". Nicholson differentiates metabolomics as the quantitative analysis of all the metabolites of an organism or biological sample, while metabonomics as the quantitative measurements of the multiparametric metabolic response of a multicellular system to pathophysiological intervention or genetic modification [67]. Nowadays, however, both metabolomics and metabonomics are usually used interchangeably.

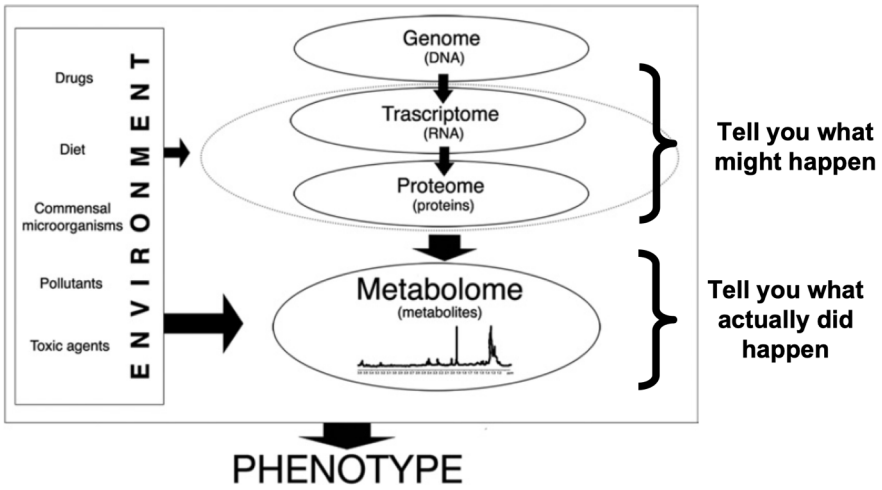

Figure 1. The Omics cascade (modified from Carraro et al., 2009) [64]

Metabolomics is a high-dimensional technology that studies the complete set of low molecular weight metabolites (<1500 Da) present at a certain time in a complex biological system, which constitute the metabolome, using body fluids such as urine, blood, or stool. Metabolites can be endogenous or exogenous, and being the downstream products of cellular function, represent a sensitive measure of the actions of upstream molecular species such as genes, transcripts, and enzymes, including the effects of disease, drugs, toxicity, and the environment so metabolomics is considered the "-omic" platform most closely related to host phenotype [64, 68, 69, 70]. Figure 2 shows the interaction of endogenous and exogenous substances in different metabolic pathways with possible overlaps.

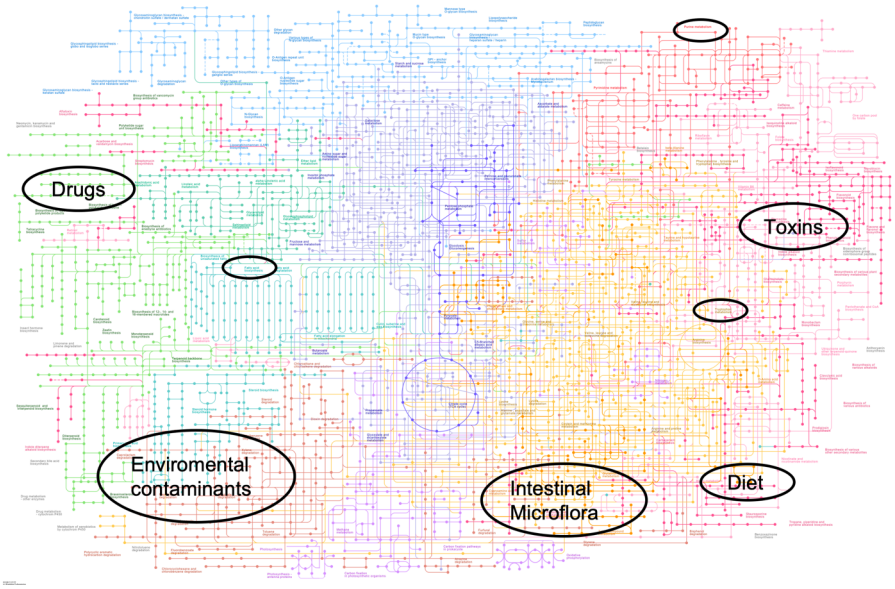

Figure 2. KEGG Metabolic pathways (modified from genome.jp) [71]

Due to the possibility to identify single or classes of metabolites, metabolomics is finding major application on several fields including medicine, pharmacology, toxicology, drug discovery and development, environmental and plant sciences [64, 72, 73].

Two approaches are mainly applied in metabolomics: untargeted and targeted.
Untargeted fingerprinting is the major technique for biomarker discovery and molecular mechanism investigation, aiming to measure all the metabolites, both known and unknown, in a sample, and thanks to identification of novel metabolites without any bias, it results ideal for hypothesis generating studies [3, 74, 75].
Untargeted approaches provide the most correct path to detect unexpected changes in metabolite concentrations. The goal is to maximize the number of metabolites detected and thus provide the opportunity to observe unexpected changes. However, a single analytical method cannot detect all metabolites in a biological system. It is therefore necessary to combine multiple analytical approaches [such as complementary chromatography methods on normal and reversed phase (C18 and HILIC)] to maximize the number of metabolites detected and improve metabolome coverage. Sample preparation in untargeted studies consists of extracting the metabolites from the biological sample in a suitable solvent for

analytical analysis. The extracted sample is analyzed with an appropriate analytical method (for example, LC-MS). The results of the mass spectrometry analysis give a chromatogram with a peaks area of each metabolite that are used as a parameter in the statistical analysis to define the semi-concentration differences between the different biological samples measured. This is called relative quantification. The biological significance of each metabolite is determined during data analysis and metabolite identification, and biological interpretation is performed at the end of the experimental workflow [76].

On the other hand, a targeted analysis, focuses on the study of a specific subset of metabolites, usually fewer than 20-30, but studies analyzing more of 500 metabolites are always more common, and are able to cover most of the pathways known and usually involved in the same pathways or they have similar functions [3, 70, 74]. Targeted metabolomics is the measurement of defined groups of chemically characterized and biochemically annotated metabolites, so is a hypothesis driven approach, useful for the validation of an untargeted analysis verifying the previous generated hypothesis [3, 74, 77].

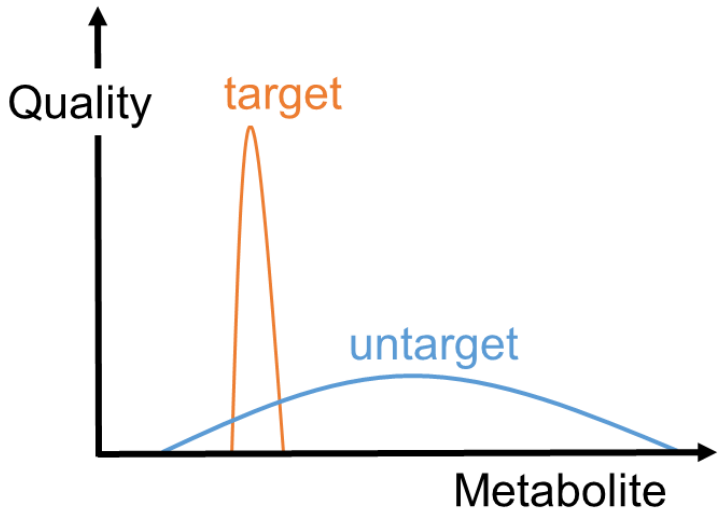

Figure 3. Number and quality of metabolites detected in targeted and untargeted MS metabolomics (modified from Gelman and Patti, 2016) [78]

Targeted methods have greater selectivity and sensitivity than untargeted methods. A targeted study can only be performed if a genuine chemical standard of the metabolite is available. The quantification of metabolites is performed using internal and chemical

standards to construct calibration curves for each of the metabolites under study and by adding internal standards during the sample preparation. Sample preparation in targeted studies applies methods that can be optimized to retain metabolites of interest and to remove other biological species and analytical artifacts that are not performed via downstream analysis. Therefore, data analysis strategies for untargeted studies require very extensive chromatogram processing. The data are produced by chemical-physical investigation techniques such as magnetic resonance spectroscopy, chromatography and mass spectrometry applied to biofluid samples or suitably selected solid tissues. These methodologies find today numerous possibilities of use in the field of medical sciences where there are numerous variables detectable on human and animal subjects that present a specific pathology. The use of metabolomic methods can help provide a holistic view of the problem, highlighting the relationships between variables and their relative importance, and can also highlight differences and similarities between samples. Considering individual biological processes as isolated processes expresses a reductionist view of vital functions, an abstraction that at times makes it possible to considerably simplify the problem under consideration but which inevitably leads to models of limited value [76, 77].

A third less used approach is the semi-targeted approach, an intermediate between untargeted and target where several metabolites, even hundreds, are identified and quantified using a single, or more, internal standard and ad hoc calibration curves for more than one analytes, providing an approximated concentration [79].

It is therefore clear that the study of metabolites can be performed with different methods: metabolite targeting is a direct approach to metabolite analysis, aiming to identify a specific metabolite, metabolite profiling identify and quantify sets of metabolites of specific metabolic pathways and metabolite fingerprinting tries to detect the metabolic characteristics that discriminate between groups of subjects, with no need to necessarily identifying every metabolite, thus being able to generate new pathophysiological hypotheses [64].

Metabolomics analysis are usually carried out on biological fluids such as blood, plasma, serum, urines and cerebrospinal, bronchoalveolar and amniotic fluid [3, 74, 80, 81, 82], whose composition could change due to the effect of environmental, pharmacological, pathological or dietary stimuli, for example, and these variations could lead to comprehend

the mechanism of action of a disease or drug therapy. Each type of sample present advantages and disadvantages, like an easily sampling and handling or pre-analytical preparation, and needs a suitable analytical technique.

Various high-througput analytical tools, such as spectroscopic (NMR, MS, MS-MS) and chromatographic (HPLC, GC, GC-MS, LC-MS, and TLC) techniques, are widely employed in metabolomics, and each approach presents its own advantages and disadvantages [83, 84].

Metabolomics studies present several steps with possible issues that need to be addressed, from sample collection to data mining, so a well-defined workflow is essential to guarantee accurate and high quality data.

3.2.1 Pre-analytical

For metabolomics studies, sampling should be performed with extreme care, due to the risk to introduce pre-analytical BIAS in samples. Each kind of sample needs a suitable sampling method and all samples must be collected, stored and treated in the same manner [80], highlighting, thus, the need of protocols to standardize the procedures.

Urine collection is scarcely invasive and relatively easy and it is possible to collect them in different and selected periods of time, based on the experiment designed. The urine samples collected are classified correctly, with the description of the collected time and day. This information is needed to study time-related trends in metabolites diurnal variation and to search biomarkers, and 24h samples, preferred to eliminate large variability in metabolite profiles obtained in shorter collection periods, when an overall status of the individual is the aim of the analysis [85].

Generally, it is important to maintain the samples at low temperature, to prevent degradation and block enzymatic reactions that would modify the metabolome.

Urine samples should be stored at -80 °C after harvesting, and cycles of sample freeze-thaw should be avoided whenever possible and the bacteriostatic preservative sodium azide should also be added [86].

Blood collection is more invasive than collecting urine, and the metabolic profiles of blood

fractions provide a different, but complementary, metabolic information compared to the ones obtained with urine [87], and small amounts, few microliters, are needed.
Blood is composed of plasma and a cell fraction and the second one should be separated by centrifugation, considering to possibly add an anticoagulant agent like heparin [88].
Infact, blood consists of two main components: plasma, an extracellular fluid containing clotting factors, proteins, glucose and minerals, and cellular elements, which are made up of blood cells and platelets. Plasma is prepared by collecting the whole blood into anticoagulant-treated tubes followed by a centrifugation step to separate blood cells. Serum is the liquid fraction of whole blood, obtained by allowing the sample to clot naturally followed by a centrifugation step and the supernatant results in being serum without cells and clotting factors. Serum samples, unlike plasma samples, do not contain the protein fibrin which is responsible for the bloods capability to clot. However, both matrices are appropriate for blood metabolomics. The metabolomics analysis of serum is known to present a higher sensitivity of metabolites compared to plasma due to the lack of big particles, but plasma has a better reproducibility due to the absence of the blood-clotting step, and the absence of platelets and the lower protein content could be beneficial to small molecule analysis, because of a reduced competition [87, 88].
As for urine samples, blood or plasma samples can be stored at -80 °C, inhibiting enzymatic activity and preserving the metabolome [87].

3.2.2 Sample preparation

Depending on the type of analysis, the samples will undergo different treatments. For example, it could be necessary to purify or dilute the samples in order to reduce the background noise, usually by dilution, solvent precipitation, solid phase or liquid phase extractions (SPE and LLE) are performed [89].
Centrifugation or filtration is also a usual step in dealing with urine samples to remove materials in suspension (e.g., calculi, cellular components or proteins). various study designs still employ high-speed centrifugation. Recently, it was found that 0.20 μm filtration is superior to centrifugation or sodium azide addition in preventing bacterial

growth during storage [85].

Direct injection or injection of dilute urine samples can be used for different purpose in metabolomics study. Injecting unmodified urines makes possible to analyze an unadulterated metabolome profiles. On the other hand, using direct injection, the sensitivity for analytes present in low concentration is hampered by the presence of co-eluting compounds and the consequent ion suppression. Also, the high salt content of urine encourages the formation of a range of adducts within the electrospray source [89, 90].

Application of sample-dilution approach, also known as ‘‘dilute and shoot’’, is only feasible when the levels of targeted metabolites are relatively high and the matrix components do not co-elute or otherwise interfere with the ionization of the analytes, thus enabling significant improvements in analytical sensitivity for the coverage of the urinary metabolome [85, 89].

The use of HPLC/UPLC-MS on the other hand, guarantees greater sensitivity for analytes present in low concentration and may be helpful in preventing the ion suppression and ion adducts [68].

For GC-MS approach, the low volatility of several metabolites (amino acids, organic acids, fatty acids) makes necessary a derivatization step prior to analyses, usually performed by silylation or methylation. Using this method, functional groups containing active hydrogen atoms (OH, NH, COOH, SH) are trimethylsilylated by a silylation reagent like MSTFA or BSTFA, thus making them suitable for GC-MS analyses [90].

NMR is a robust and reliable technique requiring minimal sample preparation, thus allowing relatively high-throughput analysis and is also non-destructive and thus preserves the biofluid and allows further analysis to be performed. The pH and the salts composition of samples, in particular of urine samples, has a significant influence on the chemical shifts in the ^{1}H-NMR spectra. Adjustment of the pH is necessary, and it is often performed by addition of a phosphate buffer at pH 7.4. When using plasma samples, it is important to consider if an anticoagulant has been added and, in case, what it is: EDTA is not suited to NMR analysis because it gives extra resonances, so the most commonly used method is to use lithium heparin tubes [91].

3.2.3 Analytical platforms

The analytical tools most used in metabolomics are liquid chromatography coupled with single-stage mass spectrometry (LC-MS) or tandem mass spectrometry (LC-MS-MS), gas chromatography combined to mass spectrometry (GC-MS), high or ultrahigh performance liquid chromatography coupled to UV or fluorescent detection (HPLC/UPLC), capillary electrophoresis (CE), and nuclear magnetic resonance (NMR) [92, 93]. The choice of the analytical techniques depends of the aim of the study and the type of samples to analyze, and each one presents both strengths and weaknesses [93].

^{1}H-NMR spectroscopy enables the detection of proton-containing metabolites in a sample, different molecules producing different signals in the NMR spectrum^{13}. NMR spectroscopy provides detailed information on molecular structure of compounds, known and unknown, both pure and in complex mixtures, especially, it is advantageous for compounds that are difficult to ionize in MS [64, 73]. It is a rapid analytical platform, characterized by high reproducibility, that requires only minimal sample handling and preparation and the NMR spectrum can generally be acquired in a few minutes. Furthermore, NMR spectroscopy is a nondestructive technique and, although it has less sensitivity than mass spectrometry (MS), it has the advantages of being relatively robust across many samples, in fact, no part of the sample becomes contaminated during the process. [88, 91, 94, 95].

A sample in a nuclear magnetic resonance (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. The substance NMR response is registered, processed, and displayed as a peak across a spectrum. The area under the peak represents the relative concentration of that metabolite compared to a reference signal, thus allowing precise quantification [3]. Absolute quantification is thus simple, it only requires an internal standard [96].

The other most used technique for metabolomics analysis is mass spectrometry MS that, as said, it is considered to have a higher sensitivity than ^{1}H-NMR [96]. MS converts the analyte molecules to a charged (ionized) state, with subsequent analysis of ions and any

fragment ions generated during the ionization process, obtaining a spectrum in which metabolites are represented according to their masses to charge ratio (m/z) [64, 97].

Electrospray ionization (ESI) is ensured by pumping the samples into a metallic capillary maintained at 3-5 kV and a stream of nitrogen nebulizes at the tip of the capillary and forms a homogeneous spray of charged droplets

The capillary is usually orthogonal to, or off-axis from, the entrance to the mass spectrometer in order to minimize the noise from neutral droplets. The droplets are rapidly evaporated by the application of heat and dry nitrogen, and the residual electrical charge on the droplets is transferred to the analytes. The intact molecular ions are produced in the ionization chamber where the ion source is kept, and then they are transferred in the mass analyzer region via several ion optics (electromagnetic elements like skimmer, focusing lens, multipole, etc.), which are basically kept to focus the ion stream to maintain a stable trajectory of the ions. The mass analyzer sorts and separates the ions according to their mass to charge ratio (m/z value). The separated ions are then passed to the detector systems to measure their concentration, and the results are displayed on a chart called a mass spectrum (see Figure 4). Since the ions in the gas phase are very reactive and often short lived, their formation and manipulation should be conducted in high vacuum. For this reason, the ion optics, analyzer, and also the detectors are kept at very high vacuum (typically from 10^{-3} torr to 10^{-6} torr pressure). Generally, the ion source is kept at atmospheric pressure, and a continuous pressure gradient and voltage gradient are used from source to the detector to help pump out the ions from source to the detector through the analyzer [97, 98].

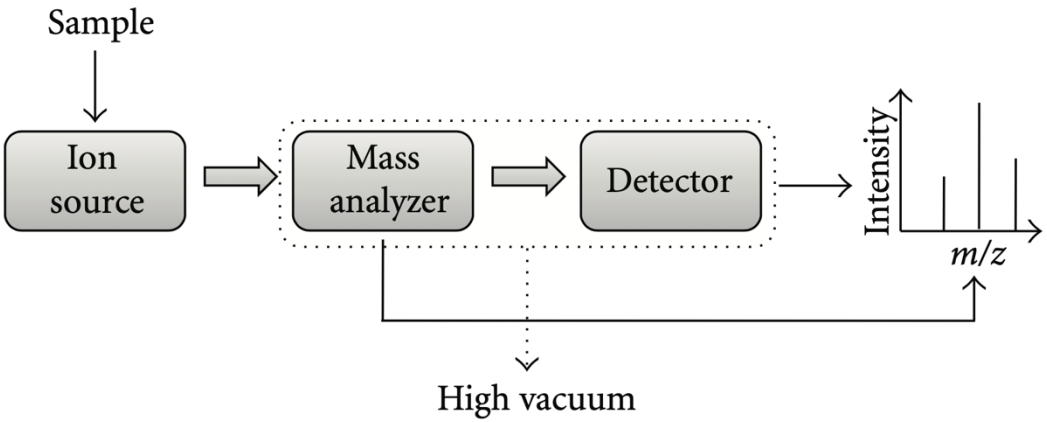

Figure 4. Basic components of an ESI MS (from Banerjee et al., 2012) [98]

Different ions sources (Electrospray Ionization Source ESI, Atmospheric Pressure Chemical Ionization Source APCI, Atmospheric pressure photo-ionization APPI) are available. ESI works well with moderately polar molecules, so it is well suited to analyze many metabolites and it is considered a "soft" ionization source, meaning that relatively little energy is imparted to the analyte, and hence little fragmentation occurs. On the other hand, APCI is useful for small molecules that are not well ionized by ESI [97].

Two different ionization modes exist, positive (POS) and negative (NEG). Positive ion mode (ESI+) is generally preferred as more compounds are expected to ionize in this mode, but in negative mode the background noise is reduced. Some metabolites are detected only in negative mode, while others were observed only in positive ion mode, so use of both positive and negative ionization offers more comprehensive metabolome coverage [99, 100].

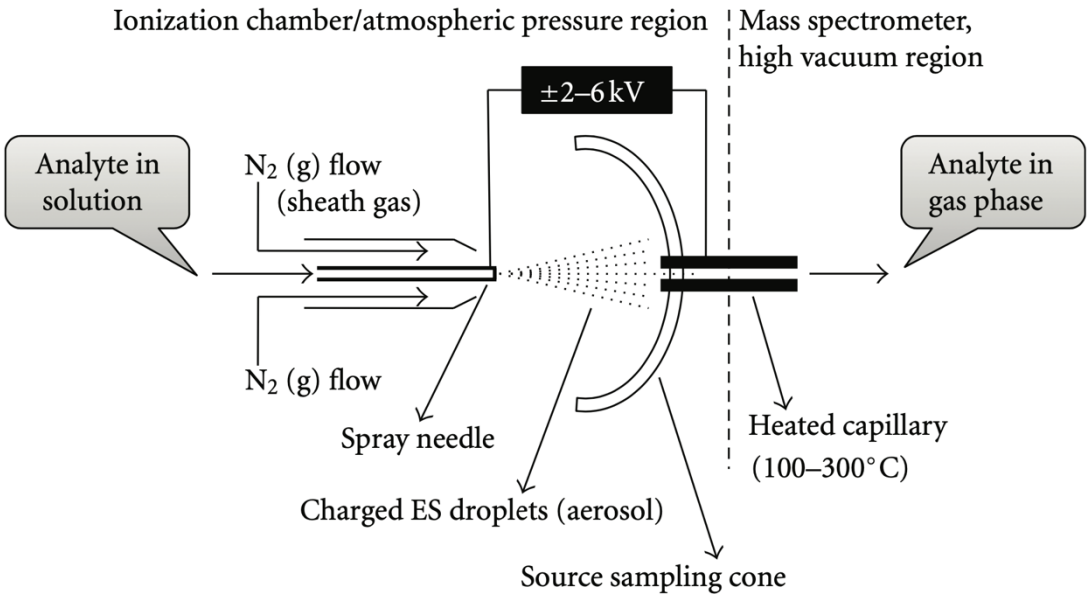

Figure 5. Representation of an ESI-ion source (from Banerjee et al., 2012) [98]

The mass analyzer is the heart of the mass spectrometer. As the choice of ion source is important, so it is the selection of the mass analyzer to use for analysis, that depends on the resolution, mass range, scan rate, and detection limit required for an application [98].

The mass analyzer can be compared with the prism. The component wavelengths of a light are separated by a prism, and then they are detected by an optical receptor. Similarly, in the mass analyzer, the different types of ions (m/z) of an ion beam are separated, and then they are passed to the detector. Magnetic (B)/electric (E) sector mass analyzer, linear quadrupole ion trap (LIT), quadrupole analyzers, orbitrap, time-of- flight mass analyzer (TOF), ion cyclotron resonance mass analyzer (ICR) and Quadrupole time-of-flight (QTOF) are some examples [97, 98, 101].

The quadrupole analyzer consists of a set of four parallel metal rods (Figure 6). A combination of constant and varying (radio frequency) voltages allows the transmission of a narrow band of m/z values along the axis of the rods. By varying the voltages with time, it is possible to scan across a range of m/z values, resulting in a mass spectrum. They usually operate at unit mass resolution meaning that the mass accuracy is seldom better than 0.1 m/z. As an alternative to scanning, the quadrupoles can be set to monitor a specific m/z value, then set to monitor another m/z value, and so on [97]. A particularly useful mass

spectrometer configuration is obtained by placing a collision cell between two quadrupole mass analyzers. This combination is called a triple quadrupole mass spectrometer and is an example of tandem MS (MS-MS) in which two or more stages of mass analysis are independently applied (Figure 6). The advantage of tandem MS is the greatly increased specificity of the analysis over single stage mass analysis. The first and third quadrupoles can also be simultaneously stepped to different m/z values, and panels of precursor/product ion pairs can be created to specifically detect a large number of targeted analytes. This process, called multiple reaction monitoring (MRM), is commonly used in LC-MS assays [97].

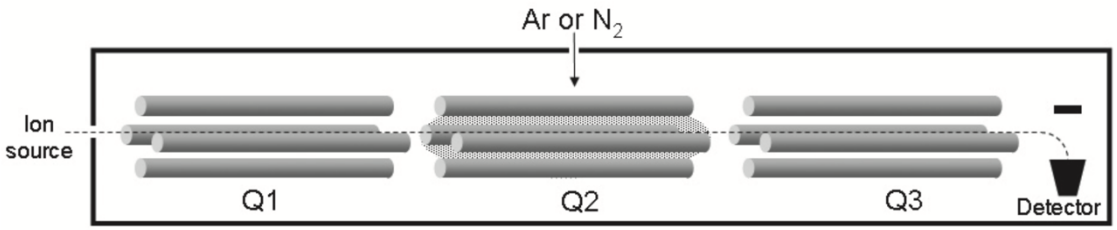

Figure 6. Representation of a triple quadrupole mass spectrometer (from Pitt, 2009) [97]. Q1 and Q3 act as mass filters and can be independently fixed, scanned or stepped. Q2 is a collision cell that contains a low pressure inert gas.

The linear time-of-flight (TOF) is the simplest mass analyzer, with a virtually unlimited mass range. The TOF reflectron, now widely used for ESI, combines time-of-flight technology with an electrostatic mirror; this offers higher resolution (typically above 5000) than a simple TOF instrument. It has gained wide use due to its fast scanning capabilities (milliseconds), good mass range (up to m/z~10000), and an accuracy in the order of 5 ppm. Quadrupole-TOF mass analyzers combine the stability of a quadrupole analyzer with the high efficiency, sensitivity, and accuracy of a time-of-flight reflectron mass analyzer and are typically coupled to electrospray ionization sources. The third quadrupole of a triple quadrupole MS can be replaced by a TOF analyzer to produce a hybrid quadrupole time-of-flight (QTOF) mass spectrometer. The quadrupole can act as a simple quadrupole analyzer to scan across a specified m/z range. However, it can also be used to selectively isolate a precursor ion and direct that ion into the collision cell. The resultant fragment ions are then analyzed by the TOF reflectron mass analyzer. Quadrupole-TOF exploits the

quadrupole's ability to select a particular ion and the ability of TOF-MS to achieve simultaneous and accurate measurements of ions across the full mass range. Quadrupole-TOF analyzers offer significantly higher sensitivity and accuracy than tandem quadrupole instruments when acquiring full-fragment mass spectra [97, 102, 103].

Figure 7 shows a scheme of a HRMS Q-TOF (Synapt G2, Waters) equipped with a quadrupole and a time-of-flight analyzer.

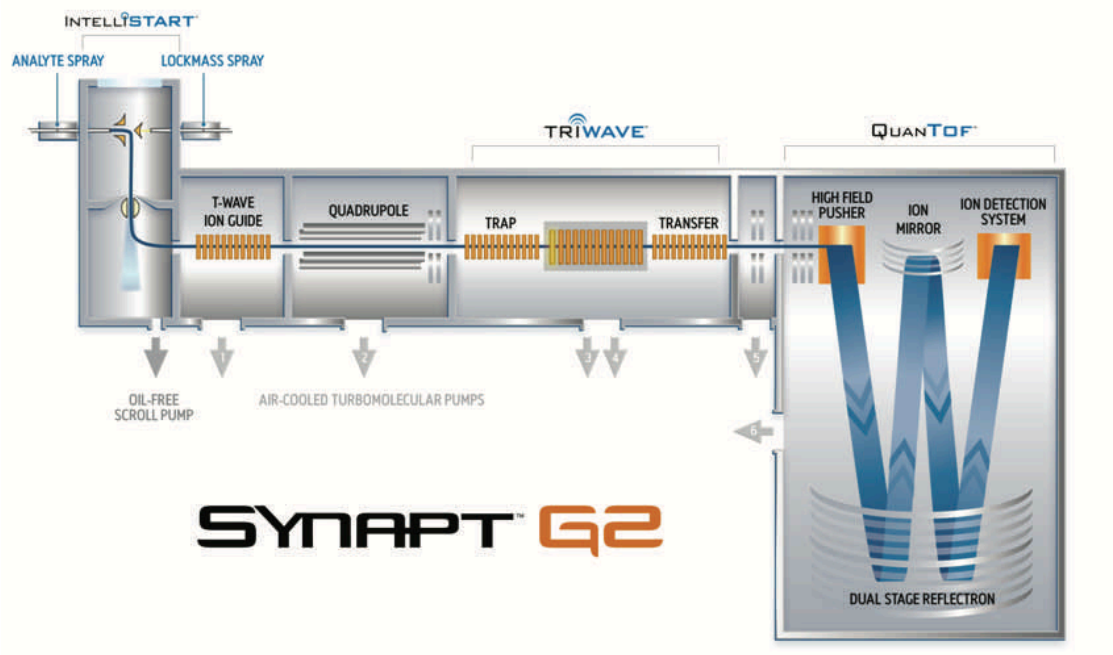

Figure 7. Schematic representation of a mass spectrometer Q-TOF (Synapt G2, Waters)

Samples can be directly injected in the instrument or, for a better compounds separation, it is possible to couple MS tools with chromatographic techniques such as gas chromatography GC, suitable for the analysis of volatile compounds, or liquid chromatography LC, useful for the study of a larger range of compounds [104].

Direct infusion-mass spectrometry (DI-MS) is an analysis platform based on the direct injection or infusion of samples into MS, without compounds separation, resulting advantageous for high-throughput screening but, although DI-MS is easy to perform, since it lacks chromatographic or electrophoretic separation capabilities, ion suppression becomes the main obstacle and isomeric compounds cannot be separated [105].

Chromatographic separation, before entering the mass spectrometer, minimizes signal suppression allowing for greater sensitivity and it can aid metabolite identification providing a retention time RT [106]. Furthermore, it possible to optimize chromatographic conditions: for LC-MS it is possible to modify mobile phases (water, methanol, acetonitrile), pH and the alternative stationary phases of chromatographic columns (normal phase, reverse-phase C8, C18, HILIC), and also MS parameters, according to the chromatographic conditions to achieve higher sensitivity in MS detection, such as capillary temperature and voltage [107, 108].

In metabolomics field, electrospray ionization mass spectrometry coupled to liquid chromatography (LC-ESI-MS) has become the most used analytical platform, providing high sensitivity and resolution, thus enabling large scale coverage of the metabolome [102, 103]. Chromatographic separation directly affects accuracy and effectiveness of the analysis, so it is necessary to use an appropriate column to achieve good separation, considering that column efficiency is determined by column length, inner diameter, mobile phase and linear velocity, and the retention time is determined by column inner diameter, film thickness and temperature [109].

Selection of appropriate column also means choosing an appropriate stationary phase in order to provide an effective chromatographic separation. Column length is a parameter to take in consideration, in fact increasing column length provides a better separation, but it also causes unwanted increase in analysis time and cost, but increasing the speed of analysis could lead inaccurate results due to co-elution of analytes [109, 110].

The coupling of ion mobility (IM) separations with LC-MS based analyses represents an emerging technology (LC-IM-MS) for metabolomics research. IM resolves gas phase ions based on their size-to-charge ratio or gas phase packing efficiency, complementing polarity and mass separations. The addition of ion mobility separation offers increased peak capacity, the ability to decrease chromatography time without sacrificing resolution, and opportunities to separate co-eluting precursors. In addition to improved mass spectra quality and increased selectivity, IM measurements can be used to determine collision cross sections (CCS) for individual metabolites. Unlike RT measurements, which vary based on column chemistry, mobile phase, and elution gradient, CCS values are physical properties and not influenced by MS or LC settings where inter-laboratory precision is reported to be at least <5% for over a broad range of molecules assayed [111].

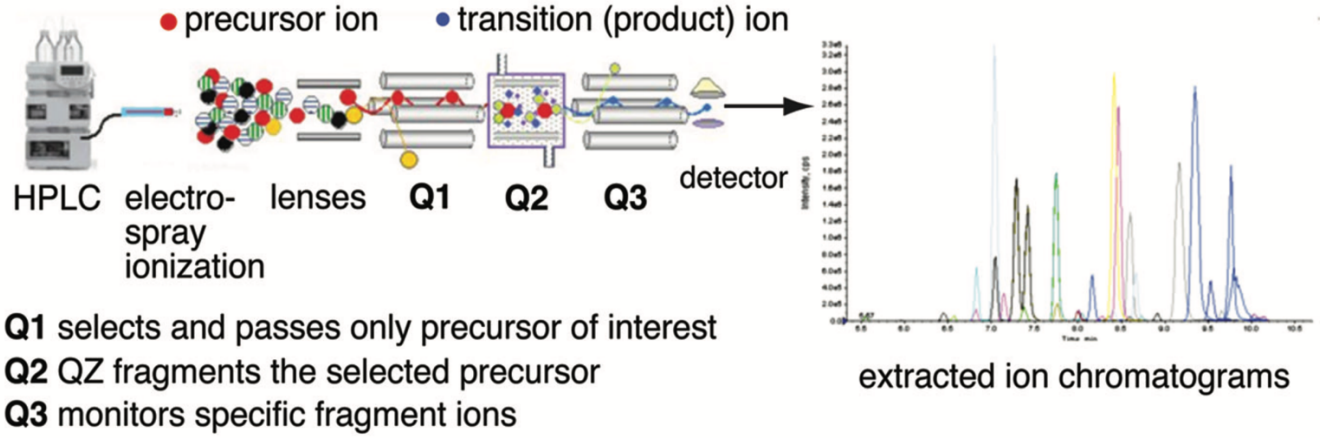

Figure 8. Representation of a LC-MS system (from Roberts et al., 2012) [77]

GC-MS is the preferred technique for volatile metabolites analysis, which yields high sensitivity and resolution, prominent reproducibility, and highly repeatable fragmentation [105]. GC-MS is a much more highly reproducible technique than LC-MS, due to the electron ionization (EI) method usually employed in which gas-phase molecules interact with kinetically activated electrons at an accepted average standard energy of 70 eV [112]. The combination of gas chromatography with electron impact ionization MS (EI-MS) provides high-chromatographic metabolite resolution, analyte-specific detection, and quantification of metabolites, as well as the capability to identify unknowns [113]. Furthermore, problems with ion suppression of co-eluting compounds are almost absent in GC-EI-MS [114].

However, many metabolites contain polar functional groups and are thermally labile at the temperatures required for their separation or are not volatile at all. Therefore, derivatization prior to GC analysis is needed [114]. To increase the volatility and thermal stability of the analytes, various derivatizations, such as alkylation, acylation, and silylation, can be employed to "protect" functional groups. Among these derivatization methods, methoximation and trimethylsilylation are commonly used in large-scale metabolomics studies with GC-MS [105]. Examples of silylation reagents are MSTFA or BSTFA, previously cited [90].

3.2.4 Data pre-processing

Raw data obtained from NMR and/or MS analysis must be treated to make them suitable for the next statistical analysis. Pre-processing involves various computational procedures where raw data from the instruments (GC/LC-MS, NMR spectra) are converted into a usable form, thus extracting all the relevant information like molecular features, for further analysis and biological interpretation [115, 116, 117, 118]. Data pre-processing for both NMR and MS includes similar procedures, such as baseline correction, peak alignment, normalization and scaling.

Distortions in the baseline can affect both chemometric analysis and quantification of metabolites. In fact, signal intensities are calculated with reference to the baseline and inadequate baseline correction can spoil the data analysis. So, baseline correction is utilized to correct the distortion in the baseline caused by systematic artifacts and, after baseline correction, spectral regions not populated with by endogenous metabolites are often removed [115, 117, 118].

Due to the nature of biological samples and variation in experimental conditions (pH, temperature, concentration, instrumental factors), peak shifts, or misalignment between identical features from different spectra, are observed. Peak alignment aims to correct this shifts for reliable identification of metabolites [115, 119].

Furthermore, for NMR binning (or bucketing) is necessary in order to reduce the data dimensionality. In binning the spectra are divided into segments (so called bins or buckets) and the total area within each bin is calculated to represent the original spectrum, in other words, it replaces the data values within each bin by a representative value. Therefore, some minor peaks shifts can be removed by spectral binning, provided that shifts of a same peak are small enough to be included in the same interval [117, 118, 119].

Then, the data are normalized, where each spectrum is multiplied by a constant term, correcting variation between samples caused by experimental sources and dilution factor, while preserving relevant biological variation. The normalization process scales the data so that different samples in a study can be compared with each other [116, 119].

Scaling, in metabolomics data analysis, refers to the column operations that are performed on each feature (spectral intensity or metabolite concentration) across all samples in order

to make the features more comparable. Commonly used scaling methods are meancentering, autoscaling and Pareto scaling and the choice of scaling technique requires a careful consideration based on the nature of the data [115, 117, 118].

Meancentering adjusts for differences between high-concentrated and low-concentrated metabolites by converting all values to vary around zero instead of around mean of metabolite level. It is usually used in combination with other scaling methods [117].

Autoscaling is commonly applied and uses the standard deviation as the scaling factor, so all metabolites have a standard deviation of one and therefore the data is analyzed on the basis of correlations instead of covariances, as is the case with centering [120].

Pareto scaling, instead of the standard deviation, uses the square root of standard deviation as scaling factor. With this method, large fold changes are decreased more than small fold changes, thus the large fold changes are less dominant compared to clean data. Furthermore, the data does not become dimensionless as after autoscaling [117, 120].

Also for MS raw data, procedures of baseline correction, peak alignment, normalization and scaling are required to make them suitable for statistical analysis. These operations are the same previously described. Furthermore, MS metabolomics data, required also noise filtering, de-isotoping, peak detection and deconvolution.

To discriminate low-intensity peaks from background noises and improve peak detection, some noise filtering methods have been developed. One is the median filtering, which removes noises by smoothing signal intensity over the course of multiple scans with a moving window of specified size, based on the fact that noises are uncorrelated from scan to scan while true signals are not. A second choice is matched filtration, which applies a filter whose coefficients are given by the second derivative of a Gaussian model peak. The adoption of such filtration can effectively reduce noise peaks whose widths are significantly less than the model peak shape [118, 121].

Since elements in nature are present in different isotopic forms, a de-isotoping step can be used to cluster the isotopic peaks corresponding to the same compounds together to simplify the data matrix by removing redundant information [115, 118].

Deconvolution extracts valuable signals and separates overlapping peaks of co-eluting components in order to improve peak quantification, but it can also introduce errors and extra variability to the process [118, 122].

There are several open access or commercial tools and websites available for data

preprocessing and analysis.

In particular, for mass spectrometry metabolomics data, some examples of software/website are MarkerLynx, Progenesis, MetAlign, MeataboAnalyst, MS-Dial, MZmine and XCMS. These tools often involve peak alignment, peak extraction, metabolites identification, and metabolic pathway analysis by searching metabolomics databases (commonly used metabolomics databases include KEGG, HMDB, Metlin, and Massbank) [105, 121].

3.2.5 Statistical analysis

MS and NMR data, after preprocessing, undergo statistical analysis to extract information. When only one variable is analyzed at a time (in omics disciplines usually one out of a panel of many measured), a so-called univariate analysis is performed. Univariate methods include tests to compare different sets of samples such as t test or ANOVA. On the other hand, when two or more variables are measured the resulting data are multivariate data [123]. So, in target metabolomics, it is possible to analyze variables with univariate methods but in untargeted analysis, due to the greater number of variables, it is not a suitable method and multivariate analysis is preferred.

Thus, most metabolomics experiments require advanced multivariate statistical methods that account for multiple experimental factors and can analyze datasets with more features than samples. These approaches can be unsupervised or supervised. The first is used to provide an overview of the study allowing the detection of trends, grouping, and possible outliers [64, 122, 124].

Principal components analysis (PCA) is an unsupervised method often used as a first approach to characterize the variation in metabolomics datasets, reduces the complexity of the data contained in the spectra and represent them by means of plots that the human eye can interpret. This approach helps to identify any intrinsic sample clustering, to see whether different groups of subjects (e.g., healthy vs ill) can be discriminated by their spectra characteristics. PCA converts the multidimensional data space into a low-dimensional model plane. This technique expresses most of the variance within a dataset using a smaller

number of factors, so called Principal Components (PCs). Each PC is a linear combination of the original variables whereby each successive PC explains the maximum amount of variance, which was not accounted for by the previous PCs. Each PC is orthogonal to the other PCs and therefore exhibits different information. The variation in spectral data is described by a few PCs, compared to the number of original variables. Conversion of the original dataset by PCA results in two matrices known as scores and loadings. In a scores plot, each point represents a single spectrum. It provides a summary of all spectra and shows how they are related to each other. Hence, the points that are close to each other have similar profiles, while objects that lie far away are characterized by different properties. The PC loadings describe the way in which the old variables are linearly combined to new variables (PCs) and indicate which variables have the greatest contribution in transforming to the new variables. In the loading plot the relation among measured variables is shown. The scores plot is mainly used to discover groups while the loadings plot is mainly used to find variables that are responsible for separating the groups. An important feature is that the directions in the score plot correspond to direction in the loading plot. Thus, any spectral clustering observed on the score plot is interpreted by examination of the loadings. [64, 117, 118, 124].

On the other hand, supervised techniques make use of a priori known structure. They use this knowledge to learn patterns and rules to predict new data supervised and to identify features of the data that distinguish experimental groups. The advantage of these methods is to provide information about those variables that indicate differences between two or more classes. Therefore, they are popular in metabolomics for biomarker discovery studies [117, 124].

Partial least squares (PLS) is a method of solving linear models and it is a latent variable approach. PLS can be used to expose relations between two matrices X and Y finding a best set of X variables that can explain most of the variation in Y. One popular method is called partial least squares discriminant analysis (PLS-DA). In PLS-DA, Y is a vector whose values represent class memberships. The recent modification of PLS-DA is orthogonal PLS-DA (OPLS- DA) in which the model is split into two parts: the systematic variations in X are split into two parts, one that is linearly related to response and one that is linearly uncorrelated to response (orthogonal). In that way only variation related to response are used to model it. In terms of prediction power OPLS-DA and PLS-DA are

comparable but, in terms of interpretability, the OPLS-DA has advantages over standard PLS-DA, since the irrelevant variation is filtered out. The scores plot shows the similarities and dissimilarities of the subjects, allowing to see that groups can be separated by the OPLS-DA model. The corresponding loadings plot can be obtained. Different from a PCA loadings plot, the metabolites identified are responsible for the classification [117, 118].
Unlike the unsupervised methods, the supervised methods enable us to predict which group a new sample belongs to on the strength of its spectra characteristics [64].

3.2.6 Metabolites identification

After the metabolite profiles are acquired at the highest possible resolution, with a high degree of separation of the chemical constituents of an extract, next crucial step is the structural identification of the single metabolites involved in the sample classification. Identifying biomarkers can involve many analytical physical and chemistry disciplines: MS and NMR have most often been used for this purpose. but no single technique fulfills all requirements for fully elucidating metabolite structure and multiple approaches need to be integrated. Due to its higher sensitivity, MS is primarily used for dereplication purposes. Fundamental support for molecular identification comes from various on-line databases. The most comprehensive available database of human metabolites is the Human Metabolome Database. Other important databases are KEGG, BioCyc, Reactome, and Metlin [64, 104].
It is also possible to identify the metabolites studying the collected spectra.

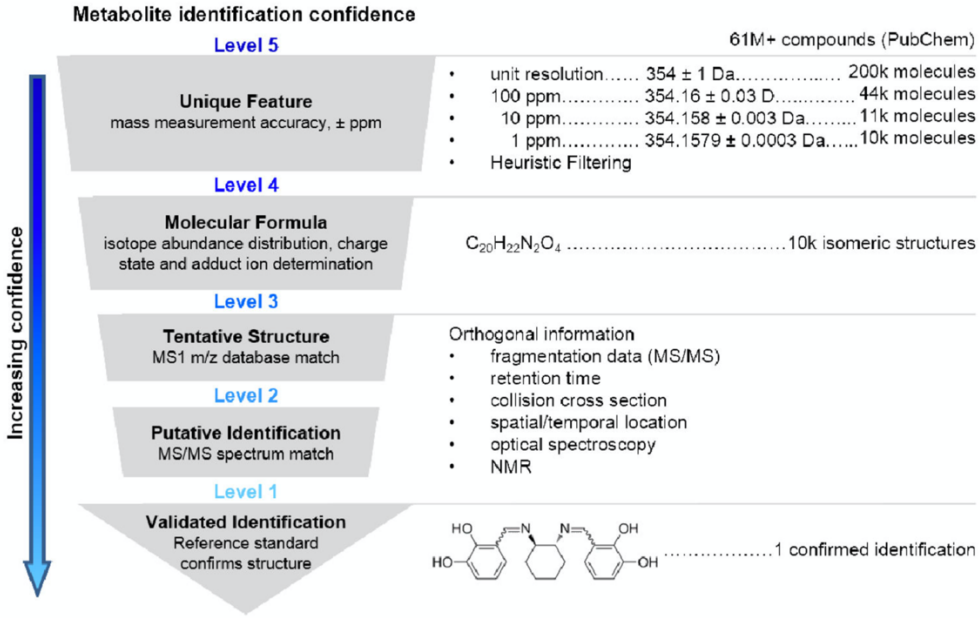

Figure 9. Workflow for metabolite identification using multidimensional MS (from Schrimpe-Rutledge et al., 2016) [111]

Because NMR chemical shifts are highly reproducible using a specific solvent or buffer conditions, comparison of the ¹H-NMR chemical shifts with those reported in databases results in unambiguous identification. For dereplication purposes, 1D proton NMR (^{1}H-NMR) spectra are used to complement the MS data obtained by metabolite profiling; this approach is often sufficient for unambiguous identification. Such complementary data can be obtained using either on-line or at-line NMR. Because ^{1}H-NMR spectroscopy detects any hydrogen-containing compound, this universal mode of detection has an advantage over other methods, which are usually limited to a specific range of compounds with particular physicochemical properties. Another advantage of NMR spectroscopy is its capacity to quantify mixtures in a non-destructive manner using adapted protocols [104].

Although 1 dimensional (1D) proton NMR spectroscopy combined with multivariate statistics is a high throughput and common approach in metabolomics studies, the problem with overlapping resonances hinders its ability to identify metabolites. 2D NMR spectroscopy offers a solution to this problem and can be used to detect and identify more metabolites than possible with the 1D approach. Different 2D NMR experiments are becoming more common in metabolomics studies with improvements in sensitivity,

reducing the acquisition times of NMR spectra. Both homonuclear and heteronuclear 2D NMR experiments have been routinely used in metabolomics analysis for signal isolation and unambiguous assignment of metabolites such as heteronuclear single quantum coherence (HSQC), that correlates proton and carbon NMR resonances and has a high degree of resolution in the second (13C) dimension, and heteronuclear multiple bond correlation (HMBC) employed to take advantage of the spectral width of carbon for metabolite discrimination and identification, where an n-bond 1H-13C shift correlation spectrum ($n > 1$) can be used to determine possible ways in which to combine the different fragments. In addition, several proton 2D homonuclear correlation spectroscopy techniques have been used in NMR-based metabolomics analyses, including total correlation spectroscopy (TOCSY), correlation spectroscopy (COSY), diffusion ordered spectroscopy (DOSY) and 2D J-resolved NMR spectroscopy [94, 117, 125].

Although the NMR approaches are capable of identifying the majority of detectable metabolites in a given sample, they can fail in the discrimination of metabolites with very similar structures and chemical shifts [126].

The data obtained by LC-MS consist in peaks without any chemical identity. Therefore, following the selection of important, discriminant features, their interpretation in a biological context is limited only by the extent to which their precise chemical identity is known. A single metabolite can produce multiple features and its identification is not always straightforward because the feature is not always observed as a protonated or deprotonated ion, but rather as adducts or fragments. Initial examination of the metabolite chromatographic data is key to determining the true monoisotopic parent in a cluster of masses detected at the same retention time RT. The identification of metabolites is labour-intensive, time-consuming and a major bottleneck in the interpretation of MS results. The gold standard for assignment of metabolites is to use internal isotope-labelled standards, or 'spiked-in' nonlabelled authentic standards. However, for untargeted metabolic profiling, the comparison of each detected metabolite with the compound standard is not feasible because of the cost and commercial availability of standards. Currently, the best approach for metabolite annotations is to use the molecular feature information (m/z, retention time and MS-MS) to determine its chemical identity against a number of freely available online databases. The potential assignments returned from the database search of a mass are then carefully examined for molecular formula and structure match, as well as the biological

relevance of the molecule to the human metabolome. Nevertheless, many of the peaks cannot be assigned to a metabolite. They are labelled as 'unknown' and are not characterized further [116].

Due to its high sensitivity, MS detection is commonly used for metabolite profiling. In most LC-HRMS applications, the metabolites are ionised using ESI that, in comparison to EI which generates fragments, has the great advantage of producing mostly molecular ion species that appear in the form of single or multiple adducts, such as $[M+H]^+$, $[M+Na]^+$ and $[M+H-H_2O]^+$ in positive ion mode or $[M-H]^-$, $[M+HCO\ ]^-$ and $[M-H+CO\ ]^-$ in negative ion mode. Furthermore, dimers, which complicate interpretation, may also be formed in the ion source. Therefore, prior to identification, the correct molecular weight MW should be defined through adduct recognition. However, definitive molecular formula determination remains a difficult task, even when mass accuracies <1 ppm are obtained, especially for compounds with a high MW (>500 Da). Feature annotation is performed by comparing an experimental mass measurement to a database of known metabolites within a mass tolerance window to generate potential candidates [104, 111].

Retention time RT is useful to identify a compound, metabolite annotation can be supported by RT predictions based upon modelling and on "projections" to similar LC methods. The major challenge regarding the use of RT for untargeted metabolomics annotation is the high number of closely eluting compounds. Furthermore, sometimes it is not stable enough to provide sufficient resolution for closely eluting compounds and even minor differences in chromatographic conditions can result in RT shifts and affect compound elution order. Consequently, using MW and RT alone for compound annotation does not always unequivocally identify a single candidate [127].

To improve annotation, additional information is required, including that provided by HR tandem MS (HR-MS-MS) and multistage MS ($HRMS^n$) spectra. The fragmentation pattern obtained by HRMS has important advantages over nominal mass measurements and fragment information generated by MS-MS or MS^n can be simply interpreted from the acquired spectra. In fact, MS-MS fragments originating from the same molecule can be used to elucidate the chemical formula and structure and therefore contribute to the discrimination of closely eluting compounds.

However, MS-MS data alone is often insufficient to differentiate structural and stereo-isomers and putative identifications require matching an experimental MS-MS spectrum

with a reference fragmentation spectrum [104, 111, 127].

LC and ion mobility can be used to generate RT and CCS information, respectively. Both of these separation methods are capable of resolving some isomeric/isobaric species [111]. So, a well-defined workflow is essential to guarantee accurate and high quality data.

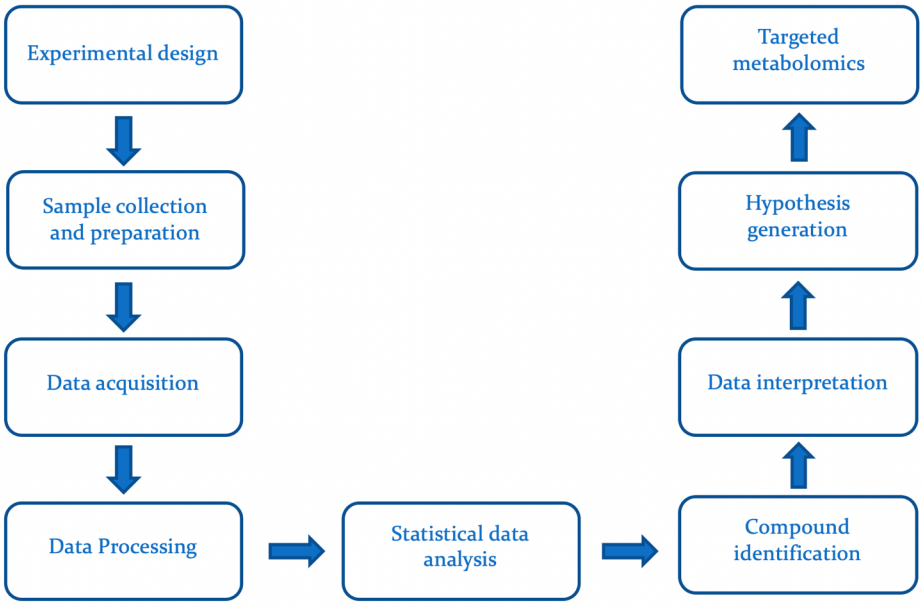

Figure 10. Metabolomics workflow

3.3 Clinical applications of metabolomics

Metabolomics, being able to detect the whole set of molecules of low molecular weight and identify metabolic patterns in cells, tissues, organs, and biological fluids, can provide a detailed phenotypic portray, representing a metabolic "snapshot", becoming an important tool in human health, allowing the exact evaluation of individual metabolic responses to pathophysiological stimuli including drugs, environmental changes, lifestyle, a great number of diseases and other epigenetics factors but also to investigate the etiology and progression of a pathological status or to formulate new hypotheses in this regard [128, 129, 130, 131].

Currently there are few clinical applications of metabolomics, but it could become a usual instrument in clinic. In particular, it could support clinical investigation of metabolic differences caused by diseases, leading to the discovery of biomarkers and biological

targets for therapeutic treatments [132].

Metabolomics resulted a highly promising tool in the early diagnosis of several fetal, perinatal, pediatric and adulthood conditions, through the detection of specific and sensible biomarkers. Moreover, metabolomics could help in monitoring the disease progression, in optimizing therapy and in the evaluation of related side effects, in the perspective of a tailored management. Innovative, interesting and promising fields of metabolomics applications are the investigation of physiological status and the diagnosis of a disease, the identification of perturbed pathways due to disease or treatment, the discovery of new and specific biomarkers, the classification of different phenotypes and functional genomics, the characterization of natural or artificial products' composition, in toxicology, pharmacology, drug development, nutrition and forensic sciences. Metabolomics also allows the study of disease phenotypes, molecular pathophysiology and cellular metabolism through metabolic profiling. Big data, provided by 'omics' tools, analyzed through the technology of machine intelligence, have also been named the black box of medicine, providing the so called "artificial intuition" [62, 128, 129, 131, 133].

3.3.1 Sepsis and metabolomics

Since sepsis is the result of pathogen's effects on the organism, metabolome alterations make metabolomics a valuable instrument in the diagnosis of sepsis, highlighting pathways changes and trying to understand the pathological mechanism [3]. In recent studies, the promising role of metabolomics in this field is reported.

GC-MS analysis of plasma samples from adult patients with sepsis versus healthy subjects detected 124 metabolites that differentiate with statistical significance between sepsis patients and control group. It was shown that the sepsis samples exhibited extensive changes in amino acids, fatty acids, and tricarboxylic acid (TCA)-cycle products, pointing out that 3 metabolic pathways (energy metabolism, amino acid metabolism, and lipid metabolism) were downregulated in sepsis patients and this may elucidate the link between organ dysfunction and host response in sepsis [134].

A meta-analysis of death-related pathways, and the following prospective validation, for the prediction of sepsis mortality using blood biomarkers differentiated metabolic profiles

between sepsis survivors and non survivors, selecting seven metabolites from the identified death-related metabolic pathways (DRMPs) including amino acids (isoleucine, alanine), lactic acid and pyruvic acid (mitochondrial metabolism), eicosanoids, and lysophospholipids, as potential biomarker for sepsis prediction [135].

Metabolomics approach is also used to predict the severity and mortality of septic shock patients.

Ferrario et al. used targeted mass spectrometry-based quantitative metabolomics approach on several series of metabolites, such as glycerophospholipids, aminoacids, biogenic amines and acylcarnitines. They suggest that alteration in kynurenine and lipid species might represent a risk factor for severe septic shock patients and a possible pathophysiological mechanism deserving further investigations [136].

A study of Liu group collected serum samples of subject at ICU admission (H0) and after 24 hours (H24). NMR analysis confirmed a previously findings of the team in investigating metabolic differences between septic shock survivors and non-survivors at H0, defining citrulline carnitine 2:0, valine, leucine, isoleucine and betanine as discriminators. Furthermore, the results showed different metabolic evolution 24h after ICU admission between septic shock survivor (SSS) and non-survivor (SSN), involving several metabolites like pyruvate, citrate, lactate, alanine and tyrosine. The time trend change in 24h show good prediction of sepsis mortality [137, 138].

In order to identify markers that distinguish between systemic inflammatory response syndrome and sepsis, Neugebauer and colleagues analyzed 406 patients determining 186 metabolites using LC-MS-MS. They find out that, in sepsis patients, serum concentration of acylcarnitines and several lipids were altered compared to systemic inflammatory response syndrome, suggesting that metabolites might be useful for differentiation and prognosis according to the type of underlying infection [139]. To identify metabolic biomarkers that could differentiate sepsis from SIRS, serum samples collected from patients with sepsis, SIRS and healthy ones have been analyzed by LC-MS-MS. The metabolic profiles of the different groups of samples was markedly different, presenting a decreased level of lactitol dehydrate and S-phenyl-D-cysteine and an increase in the levels of S-(3-methylbutanoyl)-dihydrolipoamide-E and N-nonanoyl glycine in patients with sepsis when confronted to patients with SIRS. Moreover, the profiles of patients with sepsis 48 h before death illustrated a state of metabolic disorder, with S-(3-methylbutanoyl)-

dihydrolipoamide-E, phosphatidylglycerol (22:2 (13Z, 16Z)/0:0), glycerophosphocholine and S-succinyl glutathione significantly decreased [140].

Seymour et al. determined the global metabolomic profile in plasma collected from surviving and non-surviving patients with community-acquired pneumonia (CAP) and sepsis, identifying 423 metabolites that differentiates the two groups. They demonstrated that the global metabolomic profile is different in subjects hospitalized with pneumonia and sepsis who died at 90 days compared to survivors, revealing physiologically relevant metabolites consistent with known complex processes in early sepsis, such as bile acid metabolism, protein catabolism, inflammation and oxidative stress, that were different in survivors and non-survivors, highlighting the potential for metabolomic signatures to uncover novel markers in sepsis, including putative damage-associated molecular patterns DAMPs [141].

Other metabolic differences between sepsis patients that would have survived and ones that would have not were evidenced by Langley et al. The different profiles of proteins and metabolites clustered into fatty acid transport and β-oxidation, gluconeogenesis and the citric acid cycle, and diverged more as death approached. In contrast, the metabolomes and proteomes of surviving patients with mild sepsis did not differ from survivors with severe sepsis or septic shock [142].

In prediction of bacteremic sepsis, a 6-metabolite predictive logistic model proved promising. In particular, elevated levels of myristic acid were associated with subsequent positive blood culture, demonstrating myristic acid as the most predictive metabolite in the study, and performed better than combinations of laboratory and clinical parameters. The results are encouraging because they suggest that a metabolomic approach for evaluation of patients suspected with infection can provide new diagnostic tools [143].

Recently, Elmassry et al. aimed to identify, in mice, new biomarkers after sepsis caused by *Pseudomonas aeruginosa* infection of burns. They identified, through GC-MS serum analysis, 26 metabolites that were shown to have high predictive power for *P. aeruginosa* sepsis following thermal injury, especially a combination of five metabolites (trans-4-hydroxypro- line, 5-oxoproline, glycerol-3-galactoside, indole-3-acetate, and indole-3-propionate) that could serve as a set of biomarkers for early diagnosis of sepsis caused by *P. aeruginosa* in burn patients [144].

In last decade, metabolomics has begun to be also applied in pediatric sepsis, appearing to

be promising [4].

Mickiewics et al., through NMR, analyzed serum samples of septic pediatric patients, observing that septic shock leads to significant disruption in biochemical homeostasis that strongly contributes to changes in body metabolites, allowing to distinguish between children with septic shock, those with non-infectious systemic inflammatory response syndrome and healthy children. It resulted that septic ones presented increased levels of creatinine, lactate, glucose, 2-hydroxysovalerate, 2-hydroxybutyrate and 2-oxoisocaproate and a decrease of threonine, acetate, adipate, and 2-aminobutyrate, indicating that NMR metabolite profiling might serve as a promising approach for the diagnosis and prediction of mortality in septic shock in a pediatric population [145].

Similarly, Fanos and colleagues, performed GC-MS and ^{1}H-NMR analysis on urine sample collected from LOS and EOS septic neonates, showing a compensatory reaction to a reduce level of ATP, such as an increase in glucose, ketone bodies, acetone and lactate. These biomarkers may be considered as early and reliable predictors of neonatal sepsis, suggesting that in the near future metabolomics will make possible to identify an early metabolic profile of sepsis and thus begin earlier, more targeted and effective treatments [4]. Both studies conducted by Fanos and Mickiewics groups found an increment in concentration of lactate and glucose, thus possibly suggesting a role of these metabolites in the diagnosis of sepsis. Furthermore, Mickiewics reported that combining metabolomic and inflammatory protein mediator profiling, early after symptoms presentation, may differentiate children with sepsis requiring care in a PICU from children with or without sepsis safely cared for outside a PICU, possibly aiding in making triage decisions [146].

GC-MS analysis on urine sample of a preterm neonate with sepsis fungal infection, evidenced a different metabolic profile when compared to controls, with an increase of several amino acid, like N-glycine, D-serine and L-threonine, and a decrease of hexadecanoic acid, octadecanoic acid and citric acid, suggesting an increase of the hypermetabolic state [147].

Sarafidis et al. demonstrated, through ^{1}H-NMR and LC-MS analysis of urine samples, that neonates with confirmed and suspected LOS, at the beginning of clinical manifestations, presented a different metabolic profile compared to healthy ones, thus allowing to discriminate between the two groups. Furthermore, neonates with confirmed or possible LOS exhibited comparable metabolic profiles, indicating similar metabolic alternations

upon the onset of clinical manifestations [148].

Differences in the metabolic profiles between septic shock patients with and without co-morbid conditions and healthy controls were pointed out by Pandey group, and univariate and multivariate statistical analyses were performed to identify the potential molecular biomarkers. Noted dysregulations in amino acids, carbohydrates, and lipid metabolism were observed in septic shock patients [149].

Zhang and colleagues aimed to find neonatal sepsis with meningoencephalitis-related markers using unbiased metabolomics technology and artificial intelligence analysis based on machine learning methods. The analysis of the serum and cerebrospinal fluid metabolomes combined with machine learning identified metabolite markers related to neonatal sepsis with meningoencephalitis, with changes in arginine metabolism and related changes in creatinine metabolism [150].

These results show promise in the use of metabolomics as a diagnostic tool in the diagnosis of sepsis.

4. Aim

The ultimate goal of our research program is the discovery of early biomarkers of neonatal sepsis, and in the present prospective case-control study, a two-step metabolomic approach based on mass spectrometry was applied to neonates affected by EOS in an effort to investigate the perturbations at the metabolome level that might lead to the discovery of novel early biomarkers of this condition. Firstly, an untargeted metabolomic approach was used to compare the metabolic profiles of urine samples collected within 24 h of birth from preterm neonates with and without EOS. Then, the results of this untargeted metabolomic analysis were used to guide a targeted metabolomic examination of plasma samples collected at birth from the same neonates to investigate and validate the metabolic derangements induced by sepsis further.

5. Materials and methods

5.1 Study population

This prospective, case-control study was conducted at a single, tertiary-level neonatal intensive care unit (NICU) on participants recruited from December 2015 to November 2017. Preterm neonates (<37 weeks of gestation) admitted to the NICU at the Women's and Children's Health Department of Padua Hospital (Italy) were eligible for the study. The EOS group included any neonates classified as septic infants on the basis of clinical signs (cardiovascular or respiratory instability, neurologic signs) and laboratory findings obtained within 72 h from birth (leukopenia, leukocytosis, increased C-reactive protein, increased serum lactate), in accordance with the criteria established in 2010 at an expert meeting of the European Medicines Agency on neonatal and pediatric sepsis [151]. Specifically, the criterion was the presence of at least two clinical symptoms and at least two laboratory signs of infection. These cases of the EOS group were compared with neonates who did not manifest any infection within seven days of birth, as defined by the presence of clinical and laboratory signs, in accordance with the definition of neonatal sepsis of the expert meeting of the European Medicines Agency on neonatal and pediatric sepsis [151] (controls). To avoid any influence of gestational age and weight, each neonate diagnosed with EOS was matched with the next eligible newborn of similar gestational age and weight. Infants with major congenital or chromosomal abnormalities, or with known or suspected congenital metabolic disease, asphyxiated newborns, and those given transfusions (erythrocytes, plasma, or platelets) before any collection of samples were excluded from the study. No treatment was administered before plasma sampling. At the time of urine sampling, the same pharmacological treatment, including antibiotic therapy, and the same nutrition (both parenteral and enteral) were being administered to all neonates enrolled in the study.

5.2 Samples collection

Plasma sampling: a total of 1 mL of blood collected at birth when blood tests were run on admission to the NICU (before any therapy, intravenous infusion, or milk were administered). Blood was centrifuged and the resulting plasma samples were stored at -80°C until analysis. Urine sampling: at least 2 mL of urine was collected noninvasively within 24 h of birth, by placing a cotton ball inside the newborn's nappy and checking for the presence of urine every 30 min. The cotton ball was changed if the neonate did not urinate within 3 h of its placement or if it was contaminated with fecal material. After the neonate urinated, the cotton ball was placed in the barrel of a syringe and squeezed with the plunger to collect the absorbed urine in a container prewashed with MeOH, for metabolomic analysis. The same brands of nappies and cotton balls were used throughout the study. Samples were stored at -80°C until analysis.

5.3 Metabolomic analysis

5.3.1 Metabolomic untargeted analysis

The analysis was performed at the Mass Spectrometry and Metabolomics Laboratory of the University of Padua's Women's and Children's Health Department. Urine samples were slowly thawed to ambient temperature. Each sample was stirred and centrifuged at 3600 g, then 50 μL of the supernatant from each sample were pipetted in a total recovery glass vial, adding 100 μL of 0.1% formic acid solution (dilution 1:3). Untargeted metabolic profiling was performed in positive and negative ionization mode on an Acquity Ultra Performance Liquid Chromatography (UPLC) system (Waters, U.K.) coupled to a Quadrupole Time-of-Flight (QTOF) Synapt G2 HDMS mass spectrometer (Waters MS Technologies, Ltd., Manchester, U.K.). For LC-MS analysis a Waters Acquity UPLC HSS T3 column 2.1 mm wide and 100 mm long packed with 1.8 μm beads was used and its temperature was kept at 50°C. The mobile phase flow rate was set at 0.5 ml/min. The gradient mobile phase consisted of water with 0.1% FA (A) and methanol with acetonitrile in a 90:10 ratio with

0.1% FA (B). Each sample run lasted 12 min and consisted of an isocratic phase of 5% B for 1 min, a linear increase to 30% B in 2.5 min, a linear increase to 95% B in 3 min, an isocratic phase of 95% B for 1.5 min, a washout phase of 5% B for 3 min. For each run, 5 µl of sample were injected. Mass acquisition was performed with the Quadrupole-Time-of-Flight (Q-TOF) Mass Spectrometer (Synapt G2, Waters Co.) operating at both positive-ion (ESI+) and negative-ion (ESI-) electro-spray ionization mode. The mass range scan was of 20 to 1200 amu, both in MS scan mode and in MSe mode to obtain the fragmentation spectrum of the variables that fall within the parameters set in the scan method in $MS^{\wedge E}$. The capillary voltage was set at 3.5 KV in positive mode and 2.8 KV in negative mode; the sampling cone voltage was set at 30 V in both modes. The desolvation gas flow was set at 600 L/h with temperature kept at 350 °C. The cone gas flow was set at 20 L/h with temperature kept at 110 °C. To correct for changes in environmental or experimental condition over the course of the analysis, Leucine-Enkephalin ($[M+H]^+$ = 556,2771 m/z and $[M-H]^-$ = 554,2615 m/z), at a concentration of 2 µg/ml in a solution of acetonitrile and water with 0.1% FA in a 50:50 ratio, was injected periodically (every 10 s) as internal reference (i.e. lock mass). Quality control (QC) samples and standard solution samples (Mix) were used to assess reproducibility and accuracy during the analysis, and examine the metabolite content of the samples and were injected at regular intervals throughout the sequence, together with blank samples. To further reduce analytical variability, in accordance with an in-house protocol, samples distribution in the plate and the sequence of sample injection in the UPLC-MS were randomized, and 5/6 of the fluid resulting from the addition of eluents to the sample was excluded from the ionization process (splitting). Splitting samples prevent the risk of smudge the internal surfaces of the spectrometer itself, thus reducing its sensitivity. The QCs were prepared from an aliquot of each sample, diluted with 0.1% formic acid solution with three different dilution factors (1:3, 1:5, and 1:7). The mix consisted of nine compounds of known exact mass and retention time. The QCs and mixes were injected at regular intervals during the sequence, together with blank samples to identify specific ions from the mobile phase, and any contaminants. Samples were injected randomly to prevent any spurious classification deriving from the position of the sample in the sequence.

Figure 11. Q-TOF mass spectrometer Synapt G2 (Waters) and UPLC instrumentation (Waters)

UPLC-MS data were processed by the Progenesis QI software (Waters Corporation, Milford, U.S.A.) and two data sets were generated, one for the positive-ionization mode (POS data set) and the other for the negative-ionization mode (NEG data set). The parameters used for data extraction were optimized through the preliminary analysis processing of the QC samples. We used a filter strength of 0.25 for import raw data and a QC in the middle of the sequence as a reference for the automatic alignment of all runs in the sequence. For the peak picking the sensitivity of the automatic algorithm was set at 3, where retention time limits were between 0.4 and 8 min. The so-called RT_mass variables (where RT is the retention time and mass is the mass to charge ratio m/z of the chemical compound) were generated. Variables with more than 20% of missing data were eliminated to avoid spurious statistical models generated by unrealistic combinations of the measured

variables. For each variable passing such a filter, missing data were imputed with a random number between zero and the minimum value measured for the variable. Variables with a coefficient of variation greater than 20% for QC samples have been excluded. Variables detected in the blank samples have been subtracted to the samples. The ion intensities for each peak detected were normalized, based on the calibration models obtained for the QCs with different dilution factors (1:3, 1:5, 1:7) [152]. Then, probabilistic quotient normalization was used to remove the effects of dilution on the samples' concentrations. The data tables thus generated underwent data analysis.

5.3.2 Metabolomic targeted analysis

The analysis was performed at the Mass Spectrometry and Metabolomics Laboratory of the University of Padua's Women's and Children's Health Department.

The targeted methods were developed for the metabolites revealed by metabolic profiling and implemented using UPLC coupled to a triple-quadrupole mass spectrometer.

Samples have been prepared for the analysis of amino acids, polyamines and metabolites of the kynurenine pathway.

10 µL of plasma were mixed with 10 µL polyamine internal standard (IS), 10 µL kynurenine internal standard, and 100 µL amino acids internal standard mix in methanol 0.1% v/v formic acid. The sample+IS mixtures were deproteinized and vortexed, then stored at -20 °C for 20 min and centrifuged at 13000 g for 7 min. For the analysis of the metabolites of the kynurenine pathway, 50µL of supernatant were transferred to a vial and injected to LC-MS. For amino acid and polyamine analysis, 10 µL of the supernatant were mixed in a well plate with 70 µL of borate buffer and 20 µL of AccQ-Tag reagent (AccQ-Tag Ultra Derivatization Kit, Waters Corporation), then heated for 10 min at 55°C for derivatization. The plate was placed under a stream of N_2 for 10 min to evaporate the acetonitrile of the reagent, then 20 µL of buffer or H_2O were added. 10 µL of the samples were diluted with 190 µL H_2O in another plate for high-concentration amino acid quantification.

For the analysis of neurotransmitters associated with tyrosine and tryptophan metabolism

10 µL of plasma were mixed with 5 µL of IS and 100 µL of cold acetonitrile. The sample+IS mixtures were vortexed and stored at -20 °C for 20 min, then centrifuged at 13000 g for 7 min. 10 µL of the supernatant were mixed in a well plate with 70 µL of borate buffer and 20 µL of AccQ-Tag reagent, and heated for 10 min at 55 °C for derivatization. The plate was then placed under a stream of N_2 for 10 min to evaporate the acetonitrile of the reagent, then 20 µL of buffer or H_2O were added.

Individual stock solutions in water or methanol with different percentage of formic acid depending on the different solubility of the compounds were used. A series of solution mixtures of desired concentrations were prepared by suitable dilutions of the stock solutions in 0.1% formic acid in water. All the stocks were stored at -20 °C. Stock solutions of labeled metabolites were prepared as the unlabeled and diluted as required, with water 0.1% FA, to obtain a concentration of 0.05-0.1 µM for neurotransmitters, and polyamine, and 1-10 µM for amino acids and kynurenine metabolites, and used as internal standard (IS). Calibration curves of the analytes were prepared by spiking pooled plasma, obtained from volunteers, with the diluted mixed standard solutions and IS, to the concentration ranging from 0.3 to 100 nmol/L for neurotransmitters, from 30 to 3000 nmol/L for polyamines, and from 0.05 to 250 µmol/L for amino acids and kynurenine metabolites.

Two different concentrations of QC's plasma were used for precision and accuracy. Where available we used QC from chemical companies, with 2 different level concentration (Amino Acid Quality Control set, low and TM High, Kairos Waters Corporation, Milford, MS, USA), otherwise we prepared QC by spiking pooled plasma with 2 different concentration of the analytes. The QC's plasma was extracted 2 times and analyzed 5 times within the same chromatographic run (n=10, intraday repeatability) and for 3 distinct days (n=30, between days reproducibility) to precision and reproducibility of the analytical method, expressed as coefficient of variation (CV%). Difference between measured and expected values of QC's plasma samples (Bias%) was used to estimate the accuracy of the analysis. The analytes with CV% and Bias% ≤ 20% were considered for targeted analysis. Plasma calibrations curve at 5 concentrations were built for assessing linearity, expressed as R. Sensitivity, expressed as limit of quantification (LOQ, S/N≥10), was extrapolated by lowest point of calibration solution. The R2 and LOQ for each compound were reported in Table 3.

The chemical standards and labeled standards were purchased from: Sigma-Aldrich

Corporation (Milan, Italy); Fluka (Milan, Italy); CDN Isotopes (Pointe-Claire, Quebec, Canada).; Chromsystems Instruments &Chemicals (Gräfelfing, Germany), Toronto Research Chemicals (Toronto, Ontario, Canada); Santa Cruz Biotechnology, Inc. (Dallas, Texas, USA); Coompo Research Chemicals (Wuhan, Hubei, PRC). The commercial names and the specific chemical company for each analyte were indicated in Table 3 and Table 4. The purity of all analytes and labeled internal standards was ≥98%. Water was purified with a Milli-Q Elix purification system (Millipore, Bedford, MA, USA). High-purity MS-grade solvents (formic acid, methanol, and acetonitrile) were obtained from Fluka (Milan, Italy) and used without further purification.

The analysis was conducted using a Xevo TQ-S triple-quadrupole mass spectrometer coupled to an Acquity UPLC (Waters Milford, MA, USA), interfaced with a source of Electrospray Ionization (ESI). The ESI was operated in the positive ion mode with multiple reaction monitoring (MRM). Chromatographic separation was done on a Waters Acquity UPLC HSS T3 2.1 x 100 mm 1.8 μm column (Waters Milford, MA, USA).

Figure 12. Xevo TQ-S triple-quadrupole mass spectrometer (Waters) and UPLC instrumentation (Waters)

Specific mobile phases and injection volumes were used for the different classes of metabolites as summarized below:

- Amino acids

Mobile phases consisted of waters 0.1% formic acid for phase A and acetonitrile 0.1% formic acid for phase B. Injection volume 2 µL.

Time(min)	Flow Rate	%A	%B
0	0.6	96	4
0.5	0.6	96	4

2.5	0.6	90	10
5	0.6	72	28
6	0.6	5	95
7	0.6	5	95
7.1	0.6	96	4

- Polyamines

Mobile phases consisted of waters 0.1% formic acid for phase A and acetonitrile:methanol 90:10 0.1% formic acid for phase B. Injection volume 20 µL.

Time(min)	Flow Rate	%A	%B
0	0.6	96	4
1	0.6	96	4
2.5	0.6	90	10
5	0.6	85	15
5.5	0.6	78	22
6	0.6	5	95
7	0.6	5	95
7.5	0.6	3	97
7.6	0.6	96	4
8.5	0.6	96	4

- Metabolites of the kynurenine pathway

Mobile phases consisted of waters 0.1% formic acid for phase A and acetonitrile:methanol 90:10 0.1% formic acid for phase B. Injection volume 5 μL.

Time(min)	Flow Rate	%A	%B
0	0.3	98	2
2.6	0.3	65	35
3.5	0.3	40	60
4	0.3	10	90
4.5	0.6	10	90
4.51	0.6	10	90
5.5	0.6	10	90
6	0.6	98	2
6.9	0.6	98	2
7	0.3	98	2

- Neurotransmitters associated to tyrosine and tryptophan metabolism

Mobile phases consisted of waters 0.1% formic acid for phase A and acetonitrile:methanol 90:10 0.1% formic acid for phase B. Injection volume 20 μL.

Time(min)	Flow Rate	%A	%B
0	0.6	99	1
0.5	0.6	99	1

1	0.6	96	4
3	0.6	90	10
5.5	0.6	72	28
6.5	0.6	5	95
7.5	0.6	5	95
8.01	0.6	99	1
8.5	0.6	99	1

Instrument control, data acquisition and analysis were managed with MassLynx software (version 4.1, Waters). Quantification was done using the TargetLynx function of the same software.

5.4 Statistical data analysis

The recruited neonates' demographic and perinatal data and laboratory findings were investigated applying t-test for normally distributed data, Mann-Whitney test for non-normally distributed data, and Fisher's exact test for categorical data considering statistically significant tests with p-values less than 0.05. Normality was assessed using the Shapiro-Wilk test assuming normally distributed data for p-value > 0.10. Numerical data normally distributed have been reported as mean (standard deviation), numerical data non-normally distributed as median [interquartile range], and categorical data as the number of cases (percentage) with respect to the reference group. Data emerging from untargeted and targeted metabolomics were investigated by both univariate and multivariate data analysis techniques. Untargeted data were mean-centered, whereas targeted data were auto-scaled prior to performing data analysis. Univariate data analysis was based on the t-test in the

case of normally distributed data (p-value > 0.10 for the Shapiro-Wilk test) or the Mann-Whitney test in the presence of not normally distributed data (p-value < 0.10 for the Shapiro-Wilk test), controlling the false discovery rate at the level by the Benjamini-Hochberg procedure. Multivariate data analysis was performed using projection methods. Specifically, Principal component analysis (PCA) was applied for outlier detection, while differences in the metabolic profiles of urine or plasma samples were investigated using projection to latent structures discriminant analysis (PLS-DA) with stability selection [83, 153]. Class membership was assessed by applying linear discriminant analysis to the scores of the PLS-DA model built autoscaling the dummy variables specifying the class of the sample. Because structured noise was not detected in the models, variable influence on projection score (VIP) was used for ranking the predictors within the stability selection procedure. Overall, 500 subsets were extracted by bootstrap and relevant features were identified assuming a significance level equal to 0.05. Five-fold cross-validation was applied in model optimization. Matthew correlation coefficient in calculation (MCC) and MCC calculated by five-fold cross-validation (MCC5-fold) were used to measure the goodness-of-fit and to estimate the power in the prediction of the models. Moreover, a permutation test on the class response (1000 random permutations) was performed to highlight over-fitting and to estimate the p-values of MCC and MCC5-fold. The relevant variables selected by multivariate data analysis were merged with those obtained from univariate data analysis and were annotated by searching our in-house database, the Human Metabolome Database, and the Metlin metabolite database.

Annotation for each putative marker was assigned with a different level of confidence [154], based on accurate mass, retention time, and the fragmentation patterns, where available. To improve confidence in the compound annotation, for the compounds not present in our database, theoretical fragmentation of the candidate list of compounds was performed in Progenesis, and the resulting in silico fragmentation matched against the observed fragments for a compound. Level 1 was assigned for the compounds with a difference ≤10 ppm for m/z, 0.2 min for RT, and, where available, with collision cross-section ≤2%, with respect to the standards of our in-house database, that were performed under identical analytical conditions of the current analysis. Instead, level 2 and 3 was for metabolites with m/z ≤10ppm respect to the online databases, and the fragmentation score ≥ 30 or < 30, respectively. Over-representation pathway analysis was performed

considering 80 pathways for *Homo sapiens*.

5.5 Ethical approval

The study was approved by the Ethics Committee of Padua Hospital (protocol 3636/AO/15), and written informed consent was obtained from all parents/guardians before enrolling their child.

6. Results

Among 123 eligible neonates, 15 met the criteria for early-onset sepsis (EOS group), and 15 without sepsis matched for gestational age, gender and weight were enrolled as controls. The EOS group had a gestational age of 207 (±17) days, and a birth weight of 1269 (±358) g, while the control group had a gestational age of 213 (±16) days, and a birth weight of 1300 (±354) g. The two groups did not show significant differences in gestational age (p-value = 0.34) and weight (p-value = 0.82) on the basis of the t-test. Moreover, none of the infants died, and they were all discharged home. In the EOS group, one blood culture tested positive (*Streptococcus agalactiae*); the other neonates were classified as septic infants on the basis of clinical signs (cardiovascular or respiratory instability, neurologic signs) and laboratory findings obtained within 72 h from birth (leukopenia, leukocytosis, increased C-reactive protein, increased serum lactate), in accordance with the criteria established in 2010 at an expert meeting of the European Medicines Agency on neonatal and pediatric sepsis [151]. The average time point for sepsis diagnosis was 25 h of life. Urine samples were collected within 24 h after birth for 9 cases and 10 controls. Table 1 shows the recruited neonates' demographic and perinatal characteristics, and laboratory findings at birth (no significant differences emerged between the EOS and control groups).

Table 1. The recruited neonates' demographic and perinatal characteristics, and laboratory findings at birth; numerical data normally distributed are reported as mean (standard deviation), whereas non-normally distributed data as median [interquartile range], and categorical data as the number of cases (percentage) with respect to the reference group.

Descriptive variable	EOS group (n=15)	Controls (n=15)	p
Gestational age [days]	207 (17)	213 (16)	0.34
Birth weight [g]	1269 (358)	1300 (354)	0.82
Male sex (%)	7 (47)	5 (33)	0.71
Apgar score 1 [min]	7.0 [1.5]	7 [1]	0.32
Apgar score 5 [min]	8.0 [1.5]	8 [1]	0.11
Cesarean section	14 (93)	15 (100)	>0.99
Prenatal steroids	13 (87)	13 (87)	>0.99
Small for gestational age	2 (13)	6 (40)	0.21
Positive maternal vaginal swab	3 (20)	0 (0)	0.22
Premature rupture of membranes >18 h	4 (27)	2 (13)	0.65
Inotropes	0 (0)	1 (7)	>0.99
C-reactive protein <2.9 mg/L–at birth	12 (80)	15 (100)	0.22
White blood count-day 0 [K/μL]	4.9 [2.8]	8.2 [6.2]	0.07
Platelet count-day 0 [K/μL]	200 [83]	245 [70]	0.43

While laboratory findings at birth showed no differences between the two groups, C-reactive protein value at 24–72h from birth was significantly higher in the EOS group (CRP 26 [16] mg/L) than in the control group (CRP 2.3 [1.6] mg/L) on the basis of the Mann–Whitney test (p-value < 0.001). All neonates recruited as controls did not manifest any infection within seven days of birth. In a first step, untargeted metabolic profiling of urine samples was applied to discover which are metabolic pathways perturbed by sepsis. These pathways were then analyzed more in depth in a second step using targeted methods on plasma. Two data sets were obtained in the untargeted metabolic step. Specifically, a data set including 2394 RT_mass variables was generated by the negative ionization mode (indicated as NEG data set in the following), and a data set with 3224 RT_mass variables

was obtained using the positive ionization mode (POS data set). No outliers were detected by principal component analysis (PCA) on the basis of the T2 and Q tests.

PCA has been applied to detect the presence of outliers in the untargeted metabolomic data obtained from urine samples. Specifically, T2 test and Q test have been applied with a significance level α=0.05. Data have been mean-centered and 2 principal components have been considered. In Figure 13 and Figure 14 show the score scatter plots and the T2/Q plots obtained respectively for the NEG and the POS data sets for the controls and the group of neonates developing sepsis. Red lines indicate the threshold used for outlier detection.

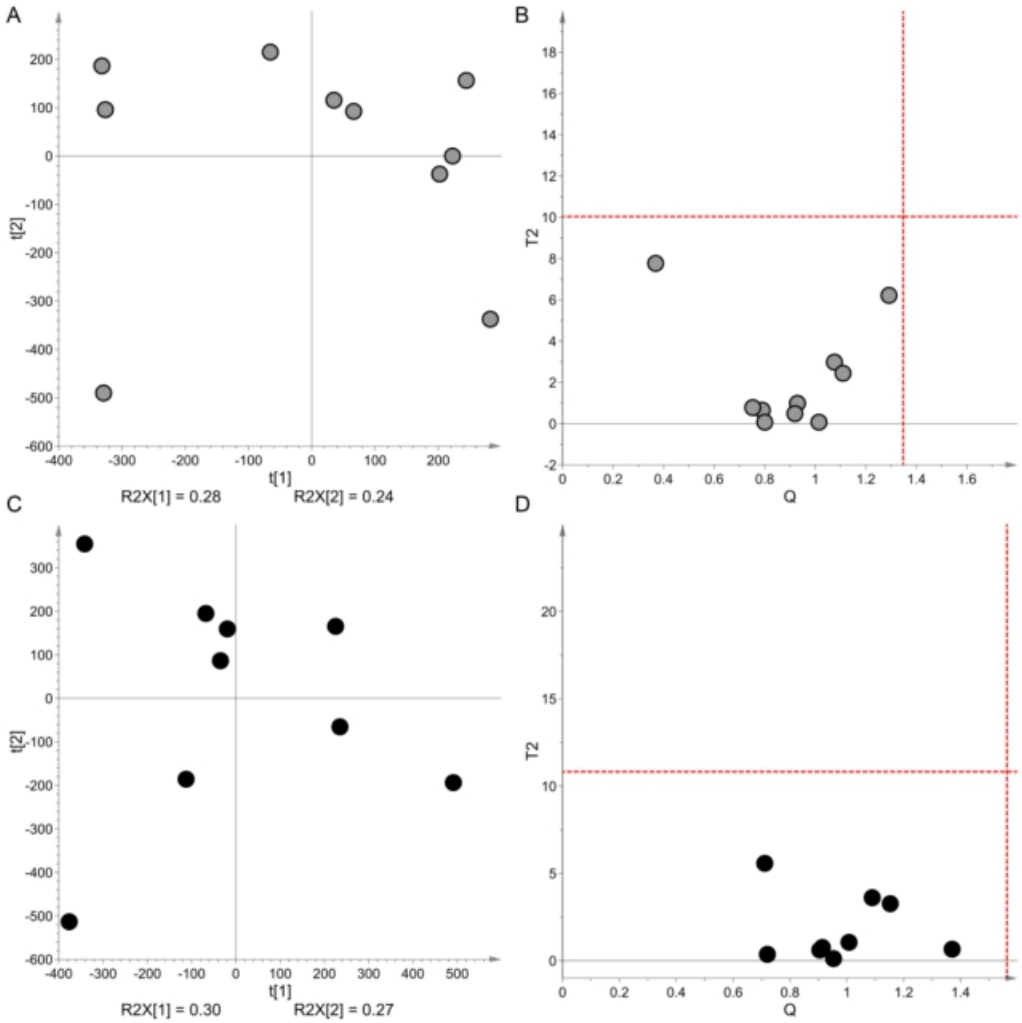

Figure 13. NEG data set: score scatter plot (panel A) and T2/Q plot (panel B) obtained for the controls and score scatter plot (panel C) and T2/Q plot (panel D) obtained for the group of neonates developing sepsis; red dashed lines indicate the threshold at the significance level α=0.05 used for outlier detection.

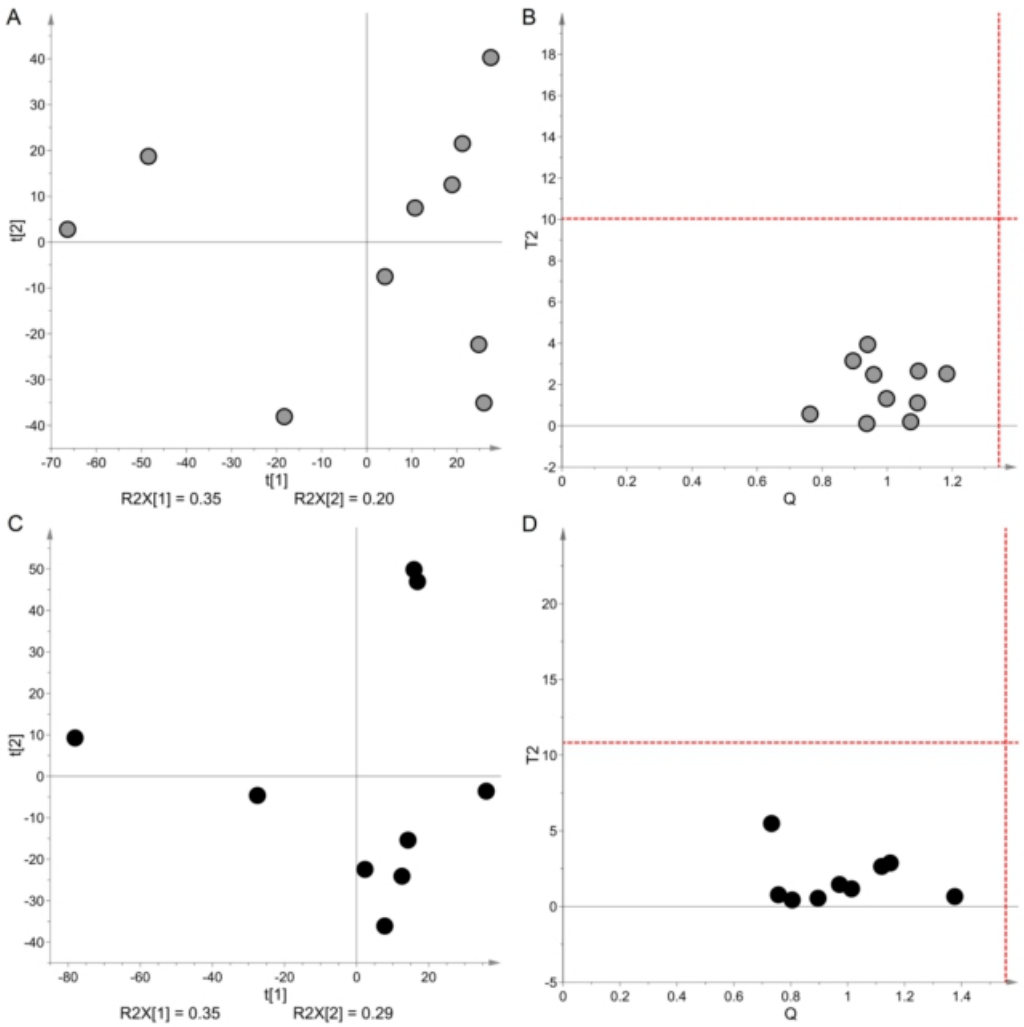

Figure 14. POS data set: score scatter plot (panel A) and T2/Q plot (panel B) obtained for the controls and score scatter plot (panel C) and T2/Q plot (panel D) obtained for the group of neonates developing sepsis; red dashed lines indicate the threshold at the significance level α=0.05 used for outlier detection.

Univariate data analysis based on t-test or Mann-Whitney test with false discovery rate highlighted 44 variables in the NEG data set, and 332 in the POS data set as relevant (δ = 0.10), for a total of 376 relevant features. For the NEG data set, projection to latent structures discriminant analysis (PLS-DA) generated a model with two latent variables, a Matthew correlation coefficient in calculation (MCC) of 0.90 (p = 0.038), and an MCC calculated by five-fold cross-validation (MCC5-fold) of 0.49 (p = 0.023). For the POS data set, a model with two latent variables, MCC = 0.90 (p-value = 0.042) and MCC5-fold = 0.37 (p-value = 0.034), were obtained. Figure 15 shows the score scatter plots obtained

with the models. It is worth noting that samples from the same group clustered in the same region of the plot, according to the high values of MCC. Stability selection has been implemented to highlight the relevant features from a multivariate point of view. All the 210 features discovered as relevant in multivariate data analysis resulted to belong to the set of 376 relevant features discovered by univariate data analysis. Thus, the relevant variables were submitted to the annotation process. Overall, 60 variables were annotated [154]: 14 at level 1 (“identified metabolites”); 44 at level 2 ("putatively annotated compounds"); and 2 at level 3 ("putatively characterized compound classes"). The annotated variables are reported in Table 2.

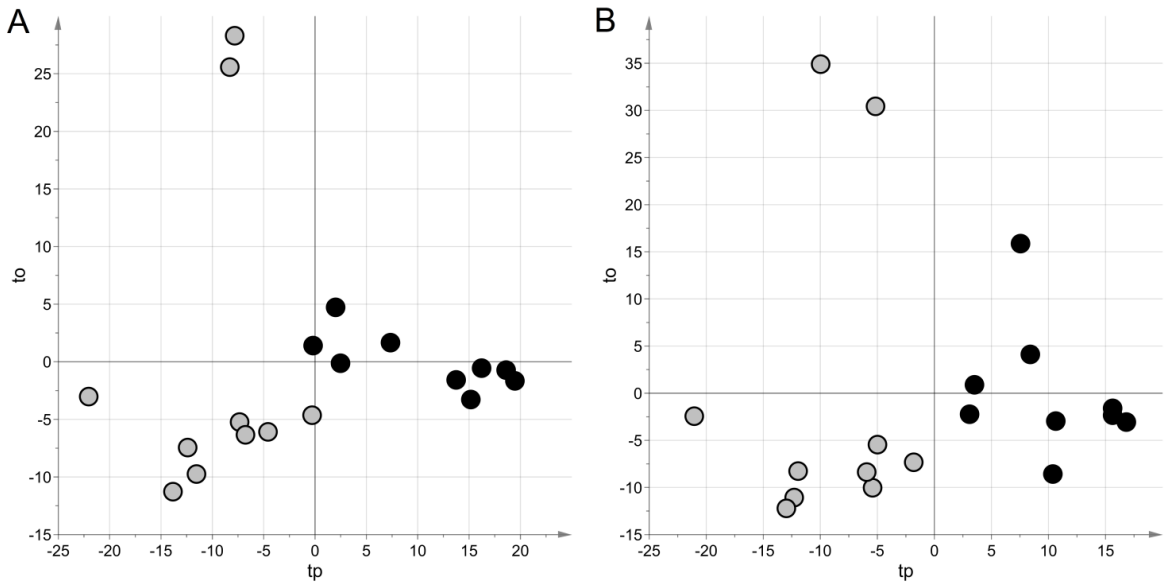

Figure 15. Untargeted metabolic profiling: score plotter plots obtained with projection to latent structures discriminant analysis (PLS-DA) modeling for urines samples. Samples from cases of sepsis are indicated with black circles, those from controls with light grey circles; panel (A) NEG data set; panel (B) POS data set. The PLS-DA models have been post-transformed to obtain the predictive latent variable tp and the non-predictive latent variable to [83].

Table 2. Annotated variables. Type: CTRL>SEPSIS indicates that the mean was higher in the control group than in the early onset sepsis (EOS) group; SEPSIS>CTRL indicates that the mean was higher in the EOS group than in the control group. HMDB ID: identifier used in the Human Metabolome Database. Level: annotation level according to reference 154. AUC (CI 95%): confidence interval at the level of 95% of the area under the receiver operating characteristic (ROC) curve.

m/z	type	HMDB ID	level	annotation	AUC (CI 95%)
1.750.244	CTRL > SEPSIS	HMDB00044	2	Ascorbic acid	0.66 - 1.00
3.280.445	CTRL > SEPSIS	HMDB00058	2	Cyclic AMP	0.53 - 1.00
1.320.772	CTRL > SEPSIS	HMDB00064	2	Creatine	0.62 - 1.00
1.300.868	CTRL > SEPSIS	HMDB00070	2	Pipecolic acid	0.58 - 1.00
1.520.575	CTRL > SEPSIS	HMDB00132	1	Guanine	0.72 - 1.00
4.643.013	CTRL > SEPSIS	HMDB00138	1	Glycocholic acid	0.66 - 1.00
1.370.464	CTRL > SEPSIS	HMDB00157	2	Hypoxanthine	0.75 - 1.00
1.800.655	CTRL > SEPSIS	HMDB00158	1	L-Tyrosine	0.60 - 1.00
1.640.708	CTRL > SEPSIS	HMDB00159	1	L-Phenylalanine	0.76 - 1.00
900.555	CTRL > SEPSIS	HMDB00161	2	L-Alanine	0.64 - 1.00
1.160.711	CTRL > SEPSIS	HMDB00162	2	L-Proline	0.65 - 1.00
1.540.618	CTRL > SEPSIS	HMDB00177	1	L-Histidine	0.60 - 1.00
1.471.132	CTRL > SEPSIS	HMDB00182	1	L-Lysine	0.54 - 1.00
239.016	CTRL > SEPSIS	HMDB00192	1	L-Cystine	0.58 - 1.00
1.650.552	CTRL > SEPSIS	HMDB00205	2	Phenylpyruvic acid	0.56 - 1.00
1.871.084	CTRL > SEPSIS	HMDB00206	1	N6-Acetyl-L-lysine	0.75 - 1.00
335.068	CTRL > SEPSIS	HMDB00229	2	Nicotinamide ribotide	0.76 - 1.00
1.260.224	CTRL > SEPSIS	HMDB00251	1	Taurine	0.70 - 1.00
2.830.677	CTRL > SEPSIS	HMDB00299	2	Xanthosine	0.70 - 1.00
1.600.607	CTRL > SEPSIS	HMDB00510	2	Aminoadipic acid	0.57 - 1.00
2.901.607	CTRL > SEPSIS	HMDB00552	2	3-Methylglutarylcarnitine	0.38 - 0.95
1.200.123	CTRL > SEPSIS	HMDB00574	1	L-Cysteine	0.61 - 1.00
1.350.309	CTRL > SEPSIS	HMDB00613	2	Erythronic acid	0.80 - 1.00
2.690.601	CTRL > SEPSIS	HMDB00676	2	L-Homocystine	0.35 - 0.94
2.090.928	CTRL > SEPSIS	HMDB00684	1	L-Kynurenine	0.61 - 1.00
1.300.867	CTRL > SEPSIS	HMDB00687	1	L-Leucine	0.65 - 1.00
1.300.506	CTRL > SEPSIS	HMDB00725	1	Hydroxyproline	0.65 - 1.00
1.880.712	CTRL > SEPSIS	HMDB00734	2	Indoleacrylic acid	0.64 - 1.00
1.920.662	CTRL > SEPSIS	HMDB00763	2	5-Hydroxyindoleacetic acid	0.59 - 1.00
1.530.413	CTRL > SEPSIS	HMDB00786	2	Oxypurinol	0.83 - 1.00

3.820.995	CTRL > SEPSIS	HMDB00912	2	Succinyladenosine	0.61 - 1.00
1.580.815	CTRL > SEPSIS	HMDB00927	2	Valerylglycine	0.66 - 1.00
2.030.818	CTRL > SEPSIS	HMDB00929	1	L-Tryptophan	0.70 - 1.00
3.851.296	CTRL > SEPSIS	HMDB00939	2	S-Adenosylhomocysteine	0.45 - 0.95
1.530.415	CTRL > SEPSIS	HMDB01182	2	6,8-Dihydroxypurine	0.85 - 1.00
136.076	CTRL > SEPSIS	HMDB01250	2	N-Acetylarylamine	0.62 - 1.00
1.500.556	CTRL > SEPSIS	HMDB01859	2	Acetaminophen	0.70 - 1.00
1.230.446	CTRL > SEPSIS	HMDB01870	2	Benzoic acid	0.60 - 1.00
860.605	CTRL > SEPSIS	HMDB02039	2	2-Pyrrolidinone	0.65 - 1.00
329.175	CTRL > SEPSIS	HMDB02121	2	Carnosol	0.65 - 1.00
1.470.444	CTRL > SEPSIS	HMDB02359	2	Phenylpropiolic acid	0.50 - 1.00
3.011.803	CTRL > SEPSIS	HMDB03955	2	19-Hydroxyandrost-4-ene-3,17-dione	0.65 - 1.00
232.028	CTRL > SEPSIS	HMDB04148	2	Dopamine 4-sulfate	0.63 - 1.00
2.570.772	CTRL > SEPSIS	HMDB04813	2	3-Methyluridine	0.52 - 1.00
6.715.588	CTRL > SEPSIS	HMDB05233	2	DG(20:1(11Z)/20:4(5Z,8Z,11Z,14Z)/0:0)[iso2]	0.60 - 1.00
2.901.356	CTRL > SEPSIS	HMDB05765	2	Ophthalmic acid	0.41 - 0.96
2.960.998	CTRL > SEPSIS	HMDB05862	2	2-Methylguanosine	0.53 - 1.00
730.301	CTRL > SEPSIS	HMDB06112	2	Malondialdehyde	0.59 - 1.00
6.092.639	SEPSIS > CTRL	HMDB00683	2	Harderoporphyrin	0.62 - 1.00
2.461.706	SEPSIS > CTRL	HMDB00688	2	Isovalerylcarnitine	0.62 - 1.00
3.050.977	SEPSIS > CTRL	HMDB01067	2	N-Acetylaspartylglutamic acid	0.50 - 0.97
2.191.111	SEPSIS > CTRL	HMDB01238	2	N-Acetylserotonin	0.58 - 1.00
2.981.126	SEPSIS > CTRL	HMDB01563	2	1-Methylguanosine	0.67 - 1.00
3.461.228	SEPSIS > CTRL	HMDB01913	2	Omeprazole	0.47 - 0.98
1.750.606	SEPSIS > CTRL	HMDB03070	2	Shikimic acid	0.67 - 1.00
3.371.276	SEPSIS > CTRL	HMDB03409	2	Berberine	0.54 - 1.00
1.790.559	SEPSIS > CTRL	HMDB03466	2	L-Gulonolactone	0.49 - 1.00
2.961.395	SEPSIS > CTRL	HMDB05037	2	Sumatriptan	0.50 - 0.99
2.751.128	SEPSIS > CTRL	Unknown	3	Unknown	0.80 - 1.00
8.192.381	SEPSIS > CTRL	Unknown	3	Unknown	0.52 - 1.00

Over-representation pathway analysis conducted on the annotated variables revealed six perturbed pathways (δ=0.15) related to: aminoacyl-tRNA biosynthesis; phenylalanine, tyrosine and tryptophan biosynthesis; nitrogen metabolism; cysteine and methionine metabolism; taurine and hypotaurine metabolism; and phenylalanine metabolism. Figure 16 summarizes the results of the over-representation pathway analysis.

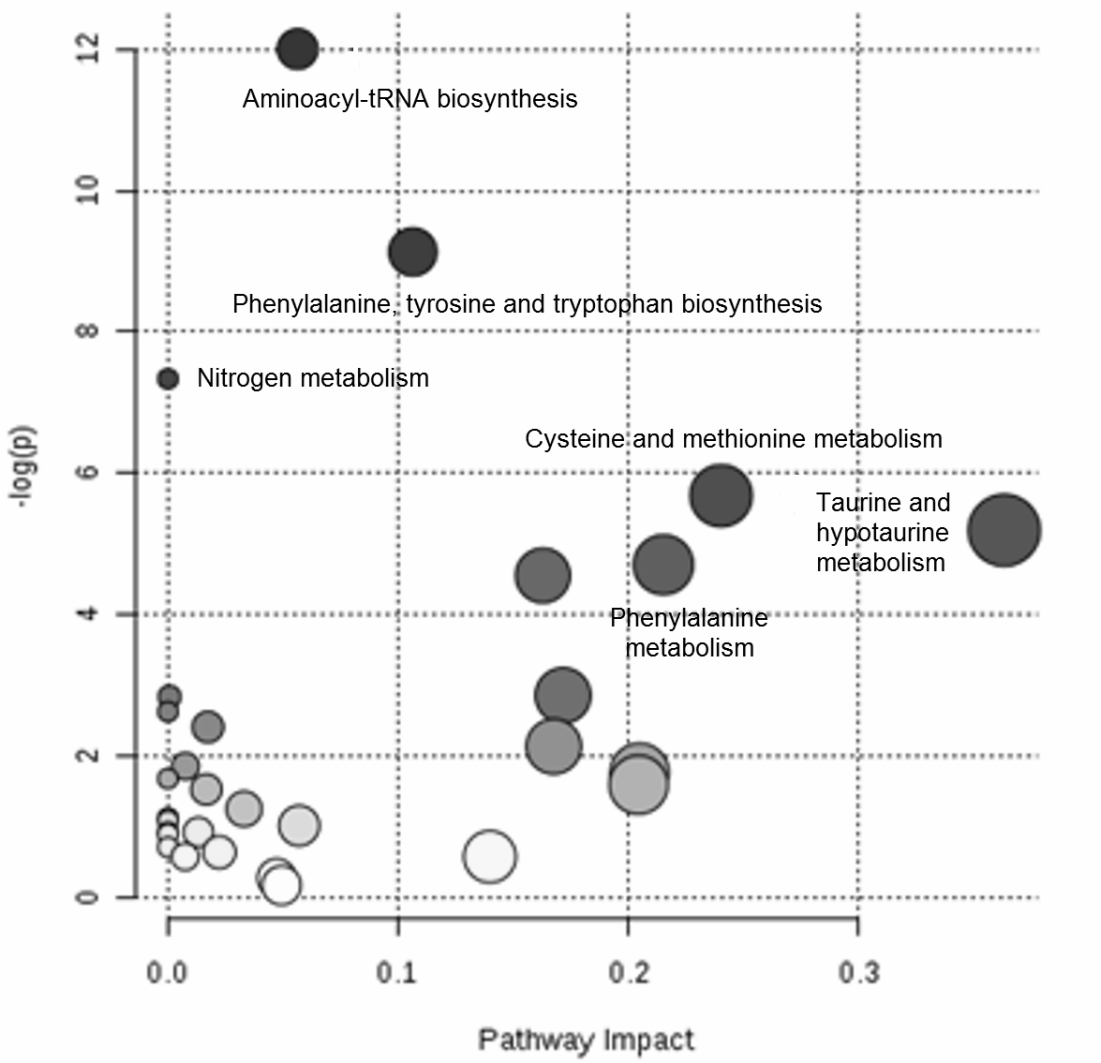

Figure 16. Untargeted metabolic profiling: over-representation pathway analysis. The impact of each perturbed pathway is shown against its negative log p-value (-log(p)). The names of the pathways with q-value less than 0.15 are shown in the figure.

In the second step of our metabolic investigation, 64 metabolites closely related to the perturbed pathways were quantified in the blood samples using targeted metabolomics to confirm the findings of the previous untargeted step. Table 3 and Table 4 report the identified and quantified metabolites by targeted methods and the labelled standards used for calibration curves.

Table 3. The name of the 64 metabolites quantified by targeted methods, their chemical group, the LOQ concentration, the value of R2 for calibration curves, and the commercial name and the companies where we purchased the standards are reported.

Metabolite	HMDB_ID	Commercial name	Company	group/pathway	LOQ	R2
2-aminobutyric acid	HMDB0000452	L-a-Amino-n-butyric Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9841
3-aminobutyric acid	HMDB0031654	3-aminobutyric acid	Sigma-Aldrich	aminoacids	2.46 µmol/L	0.9819
3-aminoisobutyric acid	HMDB0003911	D,L-ß-Aminoisobutyric Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	1
3-methylhistidine	HMDB0000479	3-Methyl-L-histidine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9999
ADMA	HMDB0001539	dimethylarginine-dihydrochloride	Sigma-Aldrich	aminoacids	0.04 µmol/L	0.9985
alanine	HMDB0000161	L-Alanine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9708
aminoadipic acid	HMDB0000510	L-a-Aminoadipic Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9909
3-aminopropanoic		ß-Alanine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9903
arginine	HMDB0000517	L-Arginine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9952
asparagine	HMDB0000168	L-Asparagine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9924
Aspartic	HMDB0000191	L-Aspartic Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9993
carnosine	HMDB0000033	L-Carnosine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9958
citrulline	HMDB0000904	L-Citrulline	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9996
cystathionine	HMDB0000099	Cystathionine*	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9987
Cystine	HMDB0000192	L-Cystine	Sigma-Aldrich	aminoacids	1.54 µmol/L	0.9976
dl-kinurenine	HMDB0000684	L-Kynurenine	Sigma-Aldrich	aminoacids/kynurenine pathways	0.12 µmol/L	0.9873
Ethanolamine	HMDB0000149	Ethanolamine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9958
GABA	HMDB0000112	g-Amino-n-butyric Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9875
glutamic acid	HMDB0000148	L-Glutamic Acid	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9974
glycine	HMDB0000123	Glycine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9918
HArg	HMDB0000670	L-homoarginine hydrochloride	Fluka	aminoacids	0.04 µmol/L	0.9985
histidine	HMDB0000177	L-Histidine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9981
homoserine	HMDB0000719	L-homoserine	Fluka	aminoacids	2.46 µmol/L	0.9922
isoleucine	HMDB0000172	L-Isoleucine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.997
leucine	HMDB0000687	L-Leucine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9995
lysine	HMDB0000182	L-Lysine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9869
methionine	HMDB0000696	L-Methionine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9988

NMMA		N-monomethyl-L-arginine	Sigma-Aldrich	aminoacids	0.04 µmol/L	0.9998
OH-lysine/allo-OH-lysine	HMDB0000450	d-DL-Hydroxylysine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.991
ornithine	HMDB0000214	L-Ornithine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9821
phenylalanine	HMDB0000159	L-Phenylalanine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.998
proline	HMDB0000162	L-Proline	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9926
sarcosine	HMDB0000271	L-Sarcosine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9987
SDMA	HMDB0003334	dimethyl-L-arginine-di(p-	Sigma-Aldrich	aminoacids	0.04 µmol/L	0.9976
serine	HMDB0000187	L-Serine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9906
taurine	HMDB0000251	Taurine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9963
threonine	HMDB0000167	L-Threonine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9988
tryptophane	HMDB0000929	L-Tryptophan	Sigma-Aldrich	aminoacids/kynurenine pathways	3.08 µmol/L	0.9943
tyrosine	HMDB0000158	L-Tyrosine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9994
valine	HMDB0000883	L-Valine	Sigma-Aldrich	aminoacids	3.08 µmol/L	0.9955
3-HAA	HMDB0001476	3-Hydroxyanthranilic acid	Sigma-Aldrich	kynurenine pathways	0.01 µmol/L	0.995
3OH-KYN	HMDB0011631	3-Hydroxy-DL-kynurenine	Sigma-Aldrich	kynurenine pathways	0.01 µmol/L	0.9974
5-HIAA	HMDB0000763	5-Hydroxyindole-3-acetic acid	Sigma-Aldrich	kynurenine pathways	0.01 µmol/L	0.9888
5-OH-ind	HMDB0001855	5-Hydroxyindole	Sigma-Aldrich	kynurenine pathways	0.05 µmol/L	0.9997
IAA	HMDB0000197	3-Indoleacetic acid	Sigma-Aldrich	kynurenine pathways	0.05 µmol/L	0.9899
IPA	HMDB0002302	Indole-3-propionic acid	Sigma-Aldrich	kynurenine pathways	0.05 µmol/L	0.9918
KYNA	HMDB0000715	kynurenic acid	Sigma-Aldrich	kynurenine pathways	0.02 µmol/L	0.9706
XA	HMDB0000881	xanthurenic acid	Sigma-Aldrich	kynurenine pathways	0.004 µmol/L	0.9719
agmatine	HMDB0001432	Agmatine sulfate salt	Sigma-Aldrich	polyamine	0.12 µmol/L	0.9931
cadaverine	HMDB0002322	Cadaverine dihydrochloride	Sigma-Aldrich	polyamine	0.09 µmol/L	0.9986
N1-AcetylSPD	HMDB0001276	N1-acetylspermidine (hydrochloride)	Cayman Chemicals	polyamine	0.03 µmol/L	0.9958
putrescine	HMDB0001414	Putrescine dihydrochloride	Sigma-Aldrich	polyamine	0.03 µmol/L	0.9922
spermidine	HMDB0001257	Spermidine trihydrocholride	Sigma-Aldrich	polyamine	0.03 µmol/L	0.9948
Spermine	HMDB0001256	Spermine tetrehydrocholride	Sigma-Aldrich	polyamine	0.06 µmol/L	0.9993
Dopamine	HMDB0000073	3-Hydroxy-Tyramine HCL	Sigma-Aldrich	neurotrasmitters	1.56 nmol/L	0.9701
epinephrine	HMDB0000068	(-) Epinephrine	Sigma-Aldrich	neurotrasmitters	1.56 nmol/L	0.9879
Metanephrine	HMDB0004063	D,L-Metanephrine Hydrochloride	Sigma-Aldrich	neurotrasmitters	1.56 nmol/L	0.9982
norepinephrine	HMDB0000216	DL-Noroadrenaline	Fluka	neurotrasmitters	2.5 nmol/L	0.9976
Octopamine	HMDB0004825	(±)-Octopamine hydrochloride	Sigma-Aldrich	neurotrasmitters	1.25 nmol/L	0.9963
Serotonin	HMDB0000259	5-Hydroxytyramine hydrochloride	Sigma-Aldrich	neurotrasmitters	0.62 µmol/L	0.9963
Synephrine	HMDB0004826	(±)-Synephrine	Sigma-Aldrich	neurotrasmitters	0.31 nmol/L	0.9918
Tryptamine	HMDB0000303	Tryptamine	Sigma-Aldrich	neurotrasmitters	3.12 nmol/L	0.9984
Tyramine	HMDB0000306	4-Hydroxyphenethylamine	Sigma-Aldrich	neurotrasmitters	3.12 nmol/L	0.9807

Table 4. The labeled standards used for calibration curves and the name of the chemical companies where we purchased the internal standards are reported.

Labeled standards	Company
Histidine D3	Sigma-Aldrich
Glycine 13C215N	Chromsystem
Arginine D7	Chromsystem
Glutamine D5	Sigma-Aldrich
Serine D3	CDN isotope
Aspartic Acid D3	Chromsystem
Citrulline D2	Chromsystem
Glutamic acid D5	Chromsystem
Alanine D4	Chromsystem
GABA D6	CDN isotope
Ornithine d6	Chromsystem
Proline D7	Sigma-Aldrich
Lysine D4	Sigma-Aldrich
Tyrosine D4	Chromsystem
Methionine D3	Chromsystem
Valine D8	Chromsystem
Leucine D3	Chromsystem
Phenylalanine D5	Chromsystem
Tryptophane D5	Sigma-Aldrich
Creatinine D3	CDN isotopes
Creatine-D3	CDN isotopes
Taurine D4	Sigma-Aldrich
5HIAA-D5	Sigma-Aldrich
L-Dopa-D3	CDN isotope
Agmatine D8	Coompo
Putrescine D4	CDN isotopes
Cadaverine D4	CDN isotopes
Spermidine 13C4	Sigma-Aldrich
Spermine D8	Sigma-Aldrich
Kynurenine-D4	Toronto Chemicals
Histamine D4	CDN isotopes
2-phenyl-d5-ethylamine (β-PEA-	CDN isotopes
Octopamine D4	CDN isotopes
Dopamine D4	CDN isotopes
Synephrine 13C215N	Santa Cruz
Norepinephrine D6	CDN isotopes
Serotonin D4	CDN isotopes
Epinephrine D6	CDN isotopes
Metanephrine D3	Sigma-Aldrich
Tryptamine D4	CDN isotopes
Serotonin D4	CDN isotopes

Specifically, 40 metabolites belonging to the families of amino acids, 9 neurotransmitters associated with tyrosine and tryptophan metabolism, 7 polyamines and 10 metabolites associated with the kynurenine pathway were quantified. Univariate data analysis identified

26 metabolites as relevant (δ=0.10).

PLS-DA generated a model with 2 latent variables, MCC=0.69 (p-value=0.08), and MCC5-fold=0.62 (p-value=0.009). Figure 17 shows the score scatter plot obtained with the model.

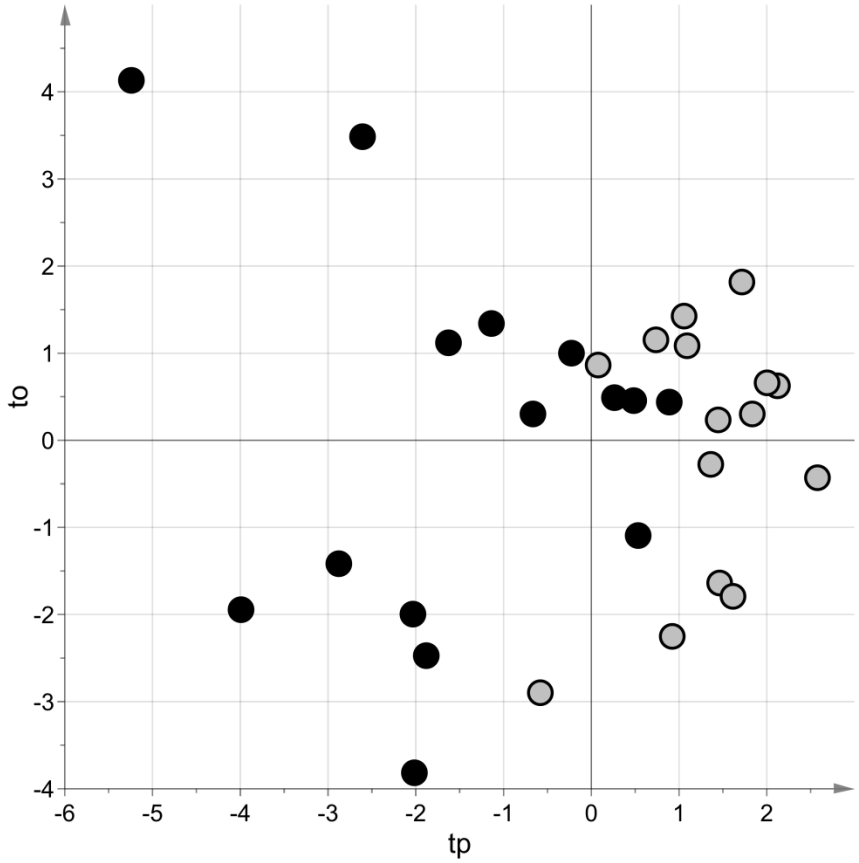

Figure 17. Targeted metabolomic investigation: score scatter plot of the PLS-DA model built using the blood samples. Samples from cases of sepsis are indicated with black circles, those from controls with light grey circles. The PLS-DA model has been post-transformed to obtain the predictive latent variable tp and the non-predictive latent variable to [83].

Stability selection led to 4 metabolites (sarcosine, ornithine, serine and cysteine) being selected as relevant for the purposes of discriminating between the groups of cases and controls. These 4 metabolites were highlighted also by univariate data analysis. Table 5 reports the relevant metabolites arising from the targeted investigation.

Table 5. Relevant metabolites for blood samples. SEPSIS>CTRL indicates that the mean in the EOS group is higher than in the control group. HMDB ID: identifier used in the Human Metabolome Database. AUC (CI 95%): confidence interval at the of 95% level of the area under the receiver operating characteristic (ROC) curve.

name	HMDB ID	type	AUC (CI 95%)
glycine	HMDB0000123	SEPSIS>CONTROL	0.59-0.96
tyrosine	HMDB0000158	SEPSIS>CONTROL	0.48-0.89
phenylalanine	HMDB0000159	SEPSIS>CONTROL	0.56-1.00
alanine	HMDB0000161	SEPSIS>CONTROL	0.50-0.92
proline	HMDB0000162	SEPSIS>CONTROL	0.52-0.95
asparagine	HMDB0000168	SEPSIS>CONTROL	0.49-0.88
lysine	HMDB0000182	SEPSIS>CONTROL	0.48-0.89
serine	HMDB0000187	SEPSIS>CONTROL	0.59-0.99
cystine	HMDB0000192	SEPSIS>CONTROL	0.52-0.92
ornithine	HMDB0000214	SEPSIS>CONTROL	0.59-0.97
serotonin	HMDB0000259	SEPSIS>CONTROL	0.49-0.88
sarcosine	HMDB0000271	SEPSIS>CONTROL	0.66-1.00
tryptamine	HMDB0000303	SEPSIS>CONTROL	0.51-0.91
tyramine	HMDB0000306	SEPSIS>CONTROL	0.51-0.91
aminoadipic acid	HMDB0000510	SEPSIS>CONTROL	0.52-0.93
kynurenine	HMDB0000684	SEPSIS>CONTROL	0.49-0.88
methionine	HMDB0000696	SEPSIS>CONTROL	0.49-0.922
kynurenic acid	HMDB0000715	SEPSIS>CONTROL	0.51-0.92
5-HIAA	HMDB0000763	SEPSIS>CONTROL	0.48-0.90
xanthurenic acid	HMDB0000881	SEPSIS>CONTROL	0.49-0.88
valine	HMDB0000883	SEPSIS>CONTROL	0.49-0.92
citrulline	HMDB0000904	SEPSIS>CONTROL	0.52-0.92
spermidine	HMDB0001257	SEPSIS>CONTROL	0.48-0.89
N1-AcetylSPD	HMDB0001276	SEPSIS>CONTROL	0.49-0.89
ADMA	HMDB0001539	SEPSIS>CONTROL	0.49-0.92
cadaverine	HMDB0002322	SEPSIS>CONTROL	0.64-0.98

Figure 18 shows the results of over-representation pathway analysis. The dysregulated pathways (δ=0.10) were associated with aminoacyl-tRNA biosynthesis, glutathione metabolism, phenylalanine, tyrosine and tryptophan biosynthesis, and tryptophan metabolism.

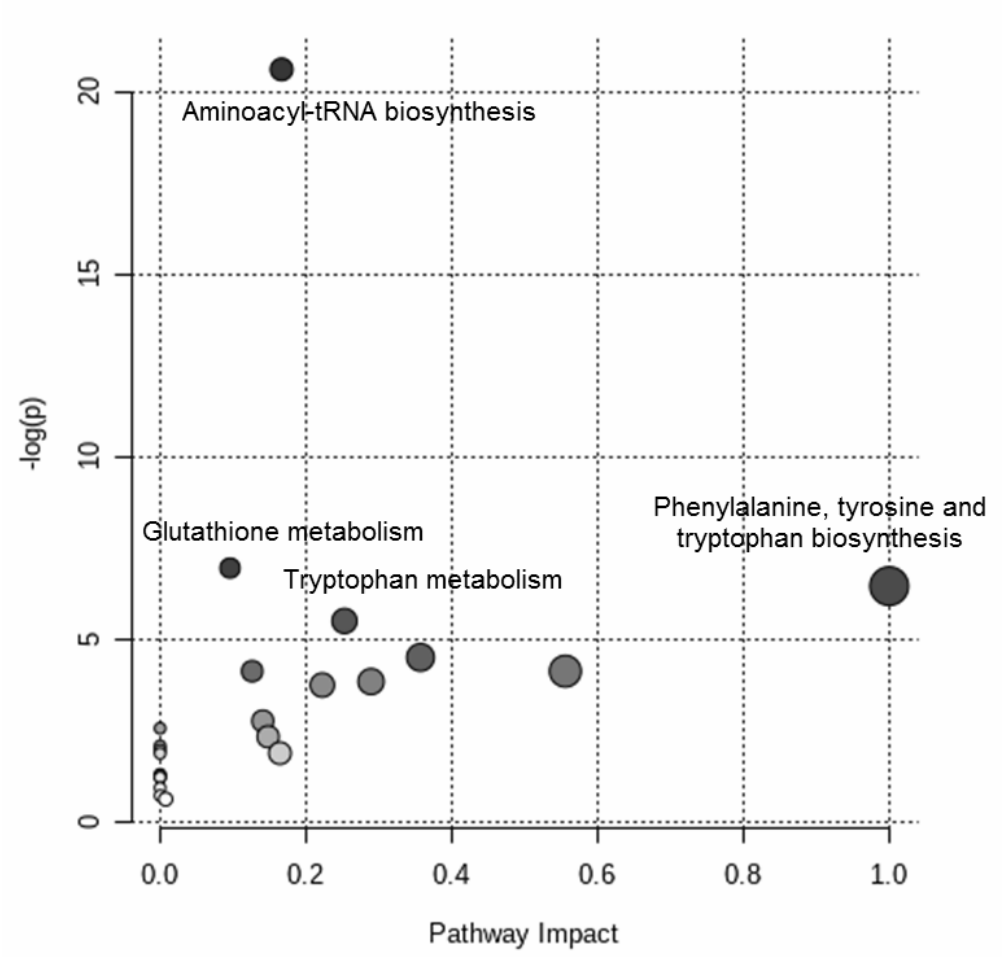

Figure 18. Targeted metabolomic investigation of the blood samples: over-representation pathway analysis. The impact of each perturbed pathway is shown against its negative log p-value (-log(p)). The names of the pathways with q-value of less than 0.10 are shown in the figure.

7. Discussion

In this study, the urine and plasma metabolome of neonates with and without EOS were examined to seek perturbations that might help identify novel early biomarkers of EOS. UPLC-MS analysis of urine samples collected within 24 h after birth revealed a clear clustering of cases of EOS compared with healthy neonates. Then, a metabolic signature exists to distinguish neonates that develop sepsis and healthy subjects. Annotating the variables derived from this untargeted analysis, putative markers discriminating between the two groups (EOS cases versus controls) were discovered (Table 2). Pathway analysis shed light on the metabolic derangements most involved in EOS (Figure 16). The metabolic pathways emerging as most significant were then further investigated using a targeted analysis on plasma samples collected at birth from the same neonates, confirming the marked disruption of the tryptophan and glutathione metabolic pathways in the neonates with EOS. Many tryptophan catabolites were enhanced in the septic neonates, and the kynurenine pathway of tryptophan metabolism was particularly stimulated. Kynurenine catabolites have both pro- and anti-oxidative properties, and they are involved in regulating glutamatergic neurotransmission and energy substrate synthesis, so tryptophan catabolism via the kynurenine pathway may have a key role in neonatal response to sepsis-induced stress [155]. Intriguingly, gut flora metabolism of tryptophan seems to be involved in metabolic anomalies induced by sepsis as well. Our results suggest that tryptophan is preferably catabolized by the gut flora of infected neonates through indole acetic acid rather than indole acrylic acid. Because the latter seems to promote the barrier function of the intestinal epithelium and mitigates inflammatory response [156, 157], a decrease in this acid could exacerbate the excessive systemic inflammatory response syndrome induced by sepsis. It has been well documented that the newborns' gut microbiome and metabolome are involved in predisposing them to sepsis and facilitating its evolution, and it can be presumed that tryptophan metabolism may be a crucial aspect of this phenomenon [158, 159, 160]. This is speculation and will be investigated in detail in future studies because the analysis of gut flora metabolism is beyond the aim of this study.

Glutathione metabolism also emerged as severely altered in neonates with EOS, in both

untargeted and targeted analyses. This is an expected finding because glutathione is a powerful antioxidant with a key role in preventing and combating reactive oxygen species [161], which are strongly generated in septic states, and the induction of its metabolism may presumably be the result of a septic process.

Interestingly, some metabolites (e.g., tyrosine, phenylalanine, kynurenine) are differently dysregulated in urine and plasma, indicating specific changes associated with the different biofluid samples. This could be due the different collection times between plasma and urine samples and also as a function of the effect of the administrated drugs present in the latter.

Metabolomics has been used widely as a diagnostic and prognostic tool for sepsis in adults. Studies on the role of metabolomics in the diagnostic work-up of pediatric and neonatal sepsis are limited, however. As explained by Mickiewics et al., we can distinguish between children with septic shock, those with non-infectious systemic inflammatory response syndrome, and healthy children from differences in their serum metabolic profiles. In addition to its diagnostic applications, the metabolomic approach has proved very accurate in establishing the prognosis for septic children [145, 146]. In a case report with two control groups, Ambroggio et al. found higher urine concentrations of metabolites previously associated with sepsis in a patient with fatal methicillin-resistant Staphylococcus aureus pneumonia than in patients with influenza pneumonia or healthy controls. These changes in urinary metabolism preceded the clinical phenotype of severe sepsis, suggesting that establishing the extent of metabolic disruption can facilitate the early identification of a sepsis phenotype ahead of its clinical diagnosis [162].

To date, only two studies have analyzed metabolic perturbations in neonatal sepsis. In the study by Fanos et al., a combined approach based on nuclear magnetic resonance and gas-chromatography/mass-spectrometry revealed a specific urinary metabolic profile in nine neonates with sepsis (both EOS and LOS) compared with 16 healthy newborns. Despite the difference in gestational age between the groups (29.1 weeks for the cases, 34.6 weeks for the controls), this was the first study to find different metabolic profiles in neonates with and without sepsis [4]. Sarafidis et al. reported that the metabolic profiles of neonates with proven or likely LOS differed considerably from those of their healthy peers. Overall, neonates with confirmed or possible LOS exhibited comparable metabolic profiles, indicating similar metabolic alternations upon the onset of clinical manifestations [148].

As far as we know, this is the first metabolomic study in which the results of untargeted

analyses on urine samples were validated using a targeted approach on plasma samples from the same neonates, which reinforces our findings. We initially analyzed urine samples because untargeted analysis requires a moderately large volume of biological fluid, and sufficient urine samples can be easily collected from preterm infants, unlike blood samples. We then confirmed the results of the first investigation by targeted analysis, which requires only a few microliters of fluid, on plasma samples collected at birth. Applying metabolomics to seeking new early biomarkers of sepsis in both urine and plasma samples offers another advantage in noninvasive urine sampling that would be particularly appropriate for neonates with only maternal risk factors or mild non-specific clinical signs of infection, who would not normally undergo blood testing. On the other hand, plasma sampling is much quicker, making it particularly useful in the sickest or most premature neonates, who would benefit most from a prompt and dependable diagnosis.

The present report also describes the first metabolomic study to focus exclusively on neonates with EOS, a condition that is a major concern for neonates born at term or preterm. Symptoms of EOS are frequently severe and take a fulminant course unless broad-spectrum antibiotic therapy is started promptly, but accurate and early biomarkers of EOS are still lacking. By revealing a clear clustering of the metabolomes of neonates with versus without EOS, our findings highlight the potential of metabolomics for discovering new early biomarkers of disease. Because the metabolism of tryptophan and glutathione was found severely disrupted in septic neonates, a few metabolites of these pathways could probably be used as biomarkers of neonatal sepsis (a hypothesis that will be investigated in future studies). Our targeted analysis revealed metabolic derangements in plasma samples obtained at birth, meaning that they occur very early on in a state of sepsis. The biomarkers emerging from this approach would therefore be very useful for the early diagnosis of EOS, which is crucial in clinical practice.

A further strength of our study lies in the low gestational age of the infants considered, which is significantly lower than in the study by Sarafidis et al. [148]. Clinical presentation is a key factor in the diagnosis of sepsis in the newborn, but the signs and symptoms are non-specific. This is particularly worrying in the case of those born very preterm, whose comorbidities may mimic sepsis. It is therefore primarily for the most premature neonates that new biochemical markers need to be discovered to help in the diagnosis of sepsis. Tools for the early diagnosis of sepsis are also much needed for preterm neonates without

any infections, who might be exposed to a greater risk of complications related to premature birth (e.g., LOS, necrotizing enterocolitis, reduced food tolerance) if given unnecessary antibiotic treatment.

A limitation of this study concerns the small number of infants recruited, but it reflects the relatively low incidence of EOS. The number of infants involved in the present study is comparable with those of previously published research using a metabolomic approach to the investigation of neonatal sepsis [4, 148].

A second potential weakness of this report is that only one newborn in the EOS group had a positive blood culture. Although blood culture is still considered the gold standard for neonatal sepsis diagnosis, the limits of this diagnostic tool are well described, especially for the diagnosis of EOS: false-negative findings are not uncommon because neonatal bacteremia is often intermittent, and intrapartum antibiotic treatment may limit the culture's diagnostic value [163, 164, 165]. To properly allocate the culture-negative neonates in the EOS group, we used the clinical and laboratory criteria identified by an expert meeting of the European Medicines Agency on neonatal and pediatric sepsis.

Finally, our results need to be confirmed, and a validation population is currently being recruited for this purpose.

Future researches will investigate, after validation of these preliminary results, the potential role of some metabolites of the tryptophan and glutathione pathways as early biomarkers of neonatal sepsis. In addition to its diagnostic accuracy in discriminating septic infants, in fact, an ideal biomarker should have many other characteristics: it should be from readily available sources, it should not be affected by comorbid conditions, and biomarker levels should vary rapidly in response to treatment, aiding in risk stratification and evaluation of prognosis. Moreover, it should be easily and rapidly measured and its analysis should not be too much expensive. If an ideal biomarker will emerge from these future researches, cost-effective methods for its measurement will be probably elaborated to fill this diagnostic gap.

8. Conclusions

In conclusion, urine and plasma samples obtained from neonates with EOS at birth showed a distinctive metabolic profile that enabled them to be clearly distinguished from those without sepsis using UPLC-MS-based analysis. The tryptophan and glutathione metabolic pathways were severely disrupted in the septic neonates. The results of this proof-of-concept study support the potential of metabolomics for elucidating the biological pathways and pathophysiological mechanisms of sepsis. Upon validation, these findings could lay the foundations for the discovery of new early biomarkers and therapeutic targets of neonatal sepsis.

Furthermore, this information could result helpful in possibly correlating the metabolic profiles with the causes of sepsis.

In future, potential metabolites of tryptophan and glutathione pathways will be tested as biomarkers of neonatal sepsis and, if one or more ideal candidates will be discovered, it will result in a greater help in the diagnosis of newborns affected by sepsis.

9. References

1. Geroulanos S, Douka ET. Historical perspective of the word "sepsis." *Intensive Care Med*. 2006;32(12):2077-2077. doi:10.1007/s00134-006-0392-2
2. Gyawali B, Ramakrishna K, Dhamoon AS. Sepsis: The evolution in definition, pathophysiology, and management. *SAGE Open Medicine*. 2019;7:205031211983504. doi:10.1177/2050312119835043
3. Evangelatos N, Bauer P, Reumann M, Satyamoorthy K, Lehrach H, Brand A. Metabolomics in Sepsis and Its Impact on Public Health. *Public Health Genomics*. 2017;20(5):274-285. doi:10.1159/000486362
4. Fanos V, Caboni P, Corsello G, et al. Urinary 1H-NMR and GC-MS metabolomics predicts early and late onset neonatal sepsis. *Early Human Development*. 2014;90:S78-S83. doi:10.1016/S0378-3782(14)70024-6
5. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*. 2016;315(8):801. doi:10.1001/jama.2016.0287
6. Rello J, Valenzuela-Sánchez F, Ruiz-Rodriguez M, Moyano S. Sepsis: A Review of Advances in Management. *Adv Ther*. 2017;34(11):2393-2411. doi:10.1007/s12325-017-0622-8
7. Fleischmann-Struzek C, Goldfarb DM, Schlattmann P, Schlapbach LJ, Reinhart K, Kissoon N. The global burden of paediatric and neonatal sepsis: a systematic review. *The Lancet Respiratory Medicine*. 2018;6(3):223-230. doi:10.1016/S2213-2600(18)30063-8
8. Vogel L. Sepsis kills one million newborns a year: WHO. *CMAJ*. 2017;189(40):E1272-E1272. doi:10.1503/cmaj.109-5504
9. Mendoza-Palomar N, Balasch-Carulla M, González-Di Lauro S, et al. Escherichia coli early-onset sepsis: trends over two decades. *Eur J Pediatr*. 2017;176(9):1227-1234. doi:10.1007/s00431-017-2975-z

10. Vaccina E, Luglio A, Ceccoli M, et al. Brief comments on three existing approaches for managing neonates at risk of early-onset sepsis. *Ital J Pediatr*. 2021;47(1):159. doi:10.1186/s13052-021-01107-3
11. Lee C-C. Changing face of neonatal early-onset sepsis: Cautious management of high-risk infants. *Pediatrics & Neonatology*. 2021;62(5):463-464. doi:10.1016/j.pedneo.2021.08.002
12. Schrag SJ, Farley MM, Petit S, et al. Epidemiology of Invasive Early-Onset Neonatal Sepsis, 2005 to 2014. *PEDIATRICS*. 2016;138(6):e20162013-e20162013. doi:10.1542/peds.2016-2013
13. Weston EJ, Pondo T, Lewis MM, et al. The Burden of Invasive Early-onset Neonatal Sepsis in the United States, 2005–2008. *Pediatric Infectious Disease Journal*. 2011;30(11):937-941. doi:10.1097/INF.0b013e318223bad2
14. Stoll BJ, Puopolo KM, Hansen NI, et al. Early-Onset Neonatal Sepsis 2015 to 2017, the Rise of *Escherichia coli* , and the Need for Novel Prevention Strategies. *JAMA Pediatr*. 2020;174(7):e200593. doi:10.1001/jamapediatrics.2020.0593
15. Fernández-Peralbo MA, Luque de Castro MD. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC Trends in Analytical Chemistry*. 2012;41:75-85. doi:10.1016/j.trac.2012.08.011
16. Klinger G, Levy I, Sirota L, et al. Outcome of Early-Onset Sepsis in a National Cohort of Very Low Birth Weight Infants. *PEDIATRICS*. 2010;125(4):e736-e740. doi:10.1542/peds.2009-2017
17. Kimpton JA, Verma A, Thakkar D, et al. Comparison of NICE Guideline CG149 and the Sepsis Risk Calculator for the Management of Early-Onset Sepsis on the Postnatal Ward. *Neonatology*. 2021;118(5):562-568. doi:10.1159/000518059
18. Good PI, Hooven TA. Evaluating Newborns at Risk for Early-Onset Sepsis. *Pediatric Clinics of North America*. 2019;66(2):321-331. doi:10.1016/j.pcl.2018.12.003
19. Stoll BJ, Hansen NI, Sánchez PJ, et al. Early onset neonatal sepsis: the burden of group B Streptococcal and E. coli disease continues [published correction appears in *Pediatrics*. 2011 Aug;128(2):390]. *Pediatrics*. 2011;127(5):817-826. doi:10.1542/peds.2010-2217

20. Dong Y, Speer CP. Late-onset neonatal sepsis: recent developments. *Arch Dis Child Fetal Neonatal Ed*. 2015;100(3):F257-F263. doi:10.1136/archdischild-2014-306213
21. Gebremedhin D, Berhe H, Gebrekirstos K. Risk Factors for Neonatal Sepsis in Public Hospitals of Mekelle City, North Ethiopia, 2015: Unmatched Case Control Study. Warburton D, ed. *PLoS ONE*. 2016;11(5):e0154798. doi:10.1371/journal.pone.0154798
22. Downey LC, Smith PB, Benjamin DK. Risk factors and prevention of late-onset sepsis in premature infants. *Early Human Development*. 2010;86(1):7-12. doi:10.1016/j.earlhumdev.2010.01.012
23. Berardi A, Sforza F, Baroni L, et al. Epidemiology and complications of late-onset sepsis: an Italian area-based study. Simeoni U, ed. *PLoS ONE*. 2019;14(11):e0225407. doi:10.1371/journal.pone.0225407
24. Gowda H, Norton R, White A, Kandasamy Y. Late-onset Neonatal Sepsis - A 10-year Review From North Queensland, Australia. *Pediatric Infectious Disease Journal*. 2017;36(9):883-888. doi:10.1097/INF.0000000000001568
25. Tsai M-H, Hsu J-F, Chu S-M, et al. Incidence, Clinical Characteristics and Risk Factors for Adverse Outcome in Neonates With Late-onset Sepsis. *Pediatric Infectious Disease Journal*. 2014;33(1):e7-e13. doi:10.1097/INF.0b013e3182a72ee0
26. Tsai M-H, Chu S-M, Lee C-W, et al. Recurrent late-onset sepsis in the neonatal intensive care unit: incidence, clinical characteristics and risk factors. *Clinical Microbiology and Infection*. 2014;20(11):O928-O935. doi:10.1111/1469-0691.12661
27. Ogundare E, Akintayo A, Aladekomo T, Adeyemi L, Ogunlesi T, Oyelami O. Presentation and outcomes of early and late onset neonatal sepsis in a Nigerian Hospital. *Afr H Sci*. 2019;19(3):2390-2399. doi:10.4314/ahs.v19i3.12
28. Adams M, Bassler D. Practice variations and rates of late onset sepsis and necrotizing enterocolitis in very preterm born infants, a review. *Transl Pediatr*. 2019;8(3):212-226. doi:10.21037/tp.2019.07.02

29. Greenberg RG, Kandefer S, Do BT, et al. Late-onset Sepsis in Extremely Premature Infants: 2000–2011. *Pediatric Infectious Disease Journal*. 2017;36(8):774-779. doi:10.1097/INF.0000000000001570
30. el Manouni el Hassani S, Berkhout DJC, Niemarkt HJ, et al. Risk Factors for Late-Onset Sepsis in Preterm Infants: A Multicenter Case-Control Study. *Neonatology*. 2019;116(1):42-51. doi:10.1159/000497781
31. Wu I-H, Tsai M-H, Lai M-Y, et al. Incidence, clinical features, and implications on outcomes of neonatal late-onset sepsis with concurrent infectious focus. *BMC Infect Dis*. 2017;17(1):465. doi:10.1186/s12879-017-2574-7
32. van Herk W, Stocker M, van Rossum AMC. Recognising early onset neonatal sepsis: an essential step in appropriate antimicrobial use. *Journal of Infection*. 2016;72:S77-S82. doi:10.1016/j.jinf.2016.04.026
33. Rehman ZU, Hassan Shah M, Afridi MNS, Sardar H, Shiraz A. Bacterial Sepsis Pathogens and Resistance Patterns in a South Asian Tertiary Care Hospital. *Cureus*. Published online May 18, 2021. doi:10.7759/cureus.15082
34. Shane AL, Sánchez PJ, Stoll BJ. Neonatal sepsis. *The Lancet*. 2017;390(10104):1770-1780. doi:10.1016/S0140-6736(17)31002-4
35. Berg D, Gerlach H. Recent advances in understanding and managing sepsis. *F1000Res*. 2018;7:1570. doi:10.12688/f1000research.15758.1
36. Procianoy RS, Silveira RC. The challenges of neonatal sepsis management. *Jornal de Pediatria*. 2020;96:80-86. doi:10.1016/j.jped.2019.10.004
37. Phua J, Ngerng W, See K, et al. Characteristics and outcomes of culture-negative versus culture-positive severe sepsis. *Crit Care*. 2013;17(5):R202. doi:10.1186/cc12896
38. Chauhan N, Tiwari S, Jain U. Potential biomarkers for effective screening of neonatal sepsis infections: An overview. *Microbial Pathogenesis*. 2017;107:234-242. doi:10.1016/j.micpath.2017.03.042
39. Faix JD. Biomarkers of sepsis. *Critical Reviews in Clinical Laboratory Sciences*. 2013;50(1):23-36. doi:10.3109/10408363.2013.764490
40. Salomão R, Ferreira BL, Salomão MC, Santos SS, Azevedo LCP, Brunialti MKC. Sepsis: evolving concepts and challenges. *Braz J Med Biol Res*. 2019;52(4):e8595. doi:10.1590/1414-431x20198595

41. Wynn JL, Polin RA. A neonatal sequential organ failure assessment score predicts mortality to late-onset sepsis in preterm very low birth weight infants. *Pediatr Res*. 2020;88(1):85-90. doi:10.1038/s41390-019-0517-2
42. Fleiss N, Coggins SA, Lewis AN, et al. Evaluation of the Neonatal Sequential Organ Failure Assessment and Mortality Risk in Preterm Infants With Late-Onset Infection. *JAMA Netw Open*. 2021;4(2):e2036518. doi:10.1001/jamanetworkopen.2020.36518
43. Iroh Tam P-Y, Bendel CM. Diagnostics for neonatal sepsis: current approaches and future directions. *Pediatr Res*. 2017;82(4):574-583. doi:10.1038/pr.2017.134
44. Huang M, Cai S, Su J. The Pathogenesis of Sepsis and Potential Therapeutic Targets. *IJMS*. 2019;20(21):5376. doi:10.3390/ijms20215376
45. Polin RA, the COMMITTEE ON FETUS AND NEWBORN, Papile L-A, et al. Management of Neonates With Suspected or Proven Early-Onset Bacterial Sepsis. *Pediatrics*. 2012;129(5):1006-1015. doi:10.1542/peds.2012-0541
46. Chiesa C, Natale F, Pascone R, et al. C reactive protein and procalcitonin: Reference intervals for preterm and term newborns during the early neonatal period. *Clinica Chimica Acta*. 2011;412(11-12):1053-1059. doi:10.1016/j.cca.2011.02.020
47. Gyllensvärd J, Ingemansson F, Hentz E, Studahl M, Elfvin A. C-reactive protein- and clinical symptoms-guided strategy in term neonates with early-onset sepsis reduced antibiotic use and hospital stay: a quality improvement initiative. *BMC Pediatr*. 2020;20(1):531. doi:10.1186/s12887-020-02426-w
48. Ershad M, Mostafa A, Dela Cruz M, Vearrier D. Neonatal Sepsis. *Curr Emerg Hosp Med Rep*. 2019;7(3):83-90. doi:10.1007/s40138-019-00188-z
49. Stocker M, van Herk W, el Helou S, et al. C-Reactive Protein, Procalcitonin, and White Blood Count to Rule Out Neonatal Early-onset Sepsis Within 36 Hours: A Secondary Analysis of the Neonatal Procalcitonin Intervention Study. *Clinical Infectious Diseases*. 2021;73(2):e383-e390. doi:10.1093/cid/ciaa876
50. Mwesigye P, Rizwan F, Alassaf N, Khan R. The Role and Validity of Diagnostic Biomarkers in Late-Onset Neonatal Sepsis. *Cureus*. Published online August 10, 2021. doi:10.7759/cureus.17065

51. Gilfillan M, Bhandari V. Neonatal sepsis biomarkers: where are we now? *RRN*. 2019;Volume 9:9-20. doi:10.2147/RRN.S163082
52. Mazzucchelli I, Garofoli F, Ciardelli L, et al. Diagnostic Performance of Triggering Receptor Expressed on Myeloid Cells-1 and CD64 Index as Markers of Sepsis in Preterm Newborns: *Pediatric Critical Care Medicine*. 2013;14(2):178-182. doi:10.1097/PCC.0b013e31826e726d
53. Abd Elkareem RM, Ahmed HM, Meabed MH, Elias SS, Elmaraghy MA. Diagnostic value of CD64 in early detection of neonatal sepsis. *Comp Clin Pathol*. 2020;29(3):639-643. doi:10.1007/s00580-020-03100
54. Delanghe JR, Speeckaert MM. Translational research and biomarkers in neonatal sepsis. *Clinica Chimica Acta*. 2015;451:46-64. doi:10.1016/j.cca.2015.01.031
55. Elmashad G, Elsayed H, Omar Z, Badr E, Omran O. Evaluation of serum amyloid A protein as a marker in neonatal sepsis. *Menoufia Med J*. 2019;32(3):1094. doi:10.4103/mmj.mmj_35_19
56. Bengnér J, Quttineh M, Gäddlin P-O, Salomonsson K, Faresjö M. Serum amyloid A – A prime candidate for identification of neonatal sepsis. *Clinical Immunology*. 2021;229:108787. doi:10.1016/j.clim.2021.108787
57. Piccioni A, Santoro MC, de Cunzo T, et al. Presepsin as Early Marker of Sepsis in Emergency Department: A Narrative Review. *Medicina*. 2021;57(8):770. doi:10.3390/medicina57080770
58. Velissaris D, Zareifopoulos N, Karamouzos V, et al. Presepsin as a Diagnostic and Prognostic Biomarker in Sepsis. *Cureus*. Published online May 13, 2021. doi:10.7759/cureus.15019
59. Saboohi E, Saeed F, Khan RN, Khan MA. Immature to total neutrophil ratio as an early indicator of early neonatal sepsis. *Pak J Med Sci*. 2019;35(1). doi:10.12669/pjms.35.1.99
60. Lavrik IN, Zhivotovsky B. Systems biology: a way to make complex problems more understandable. *Cell Death Dis*. 2014;5(5):e1256-e1256. doi:10.1038/cddis.2014.195
61. Debnath M, Prasad GBKS, Bisen PS. Omics Technology. In: *Molecular Diagnostics: Promises and Possibilities*. Springer Netherlands; 2010:11-31. doi:10.1007/978-90-481-3261-4_2

62. Akçan R, Taşteki˙N B, Yildirim MŞ, Aydogan HC, Sağlam N. Omics era in forensic medicine: towards a new age. *Turk J Med Sci*.:11.
63. Weissenbach J. The rise of genomics. *Comptes Rendus Biologies*. 2016;339(7-8):231-239. doi:10.1016/j.crvi.2016.05.002
64. Carraro S, Giordano G, Reniero F, Perilongo G, Baraldi E. Metabolomics: A New Frontier for Research in Pediatrics. *The Journal of Pediatrics*. 2009;154(5):638-644. doi:10.1016/j.jpeds.2009.01.014
65. Dopazo J. Genomics and transcriptomics in drug discovery. *Drug Discovery Today*. 2014;19(2):126-132. doi:10.1016/j.drudis.2013.06.003
66. Alseekh S, Fernie AR. Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant J*. 2018;94(6):933-942. doi:10.1111/tpj.13950
67. Nicholson JK. Global systems biology, personalized medicine and molecular epidemiology. *Mol Syst Biol*. 2006;2(1):52. doi:10.1038/msb4100095
68. Nordström A, O'Maille G, Qin C, Siuzdak G. Nonlinear Data Alignment for UPLC−MS and HPLC−MS Based Metabolomics: Quantitative Analysis of Endogenous and Exogenous Metabolites in Human Serum. *Anal Chem*. 2006;78(10):3289-3295. doi:10.1021/ac060245f
69. Wishart DS. Current Progress in computational metabolomics. *Briefings in Bioinformatics*. 2007;8(5):279-293. doi:10.1093/bib/bbm030
70. Tan SZ, Begley P, Mullard G, Hollywood KA, Bishop PN. Introduction to metabolomics and its applications in ophthalmology. *Eye*. 2016;30(6):773-783. doi:10.1038/eye.2016.37
71. KEGG pathway: Metabolic Pathways - Reference pathway - genome. KEGG: Kyoto Encyclopedia of Genes and Genomes. https://www.genome.jp/pathway/map01100. Accessed March 9, 2022.
72. Wishart DS. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat Rev Drug Discov*. 2016;15(7):473-484. doi:10.1038/nrd.2016.32
73. Salem MA, Perez de Souza L, Serag A, et al. Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. *Metabolites*. 2020;10(1):37. doi:10.3390/metabo10010037

74. Kennedy AD, Wittmann BM, Evans AM, et al. Metabolomics in the clinic: A review of the shared and unique features of untargeted metabolomics for clinical research and clinical testing. *J Mass Spectrom*. 2018;53(11):1143-1154. doi:10.1002/jms.4292
75. Che N, Ma Y, Ruan H, et al. Integrated semi-targeted metabolomics analysis reveals distinct metabolic dysregulation in pleural effusion caused by tuberculosis and malignancy. *Clinica Chimica Acta*. 2018;477:81-88. doi:10.1016/j.cca.2017.12.003
76. Lelli V, Belardo A, Timperio AM. From Targeted Quantification to Untargeted Metabolomics. 2021; doi: http://dx.doi.org/10.5772/intechopen.96852
77. Roberts LD, Souza AL, Gerszten RE, Clish CB. Targeted Metabolomics. *Current Protocols in Molecular Biology*. 2012;98(1). doi:10.1002/0471142727.mb3002s98
78. Gelman SJ, Patti GJ. Profiling cancer metabolism at the 'omic' level: a last resort or the next frontier? *Cancer Metab*. 2016;4(1):2. doi:10.1186/s40170-016-0144-x
79. Belhaj MR, Lawler NG, Hoffman NJ. Metabolomics and Lipidomics: Expanding the Molecular Landscape of Exercise Biology. *Metabolites*. 2021;11(3):151. doi:10.3390/metabo11030151
80. Turi KN, Romick-Rosendale L, Ryckman KK, Hartert TV. A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma. *Journal of Allergy and Clinical Immunology*. 2018;141(4):1191-1201. doi:10.1016/j.jaci.2017.04.021
81. Salihovic S, Broeckling CD, Ganna A, et al. Non-targeted urine metabolomics and associations with prevalent and incident type 2 diabetes. *Sci Rep*. 2020;10(1):16474. doi:10.1038/s41598-020-72456-y
82. Fotiou M, Fotakis C, Tsakoumaki F, et al. 1H NMR-based metabolomics reveals the effect of maternal habitual dietary patterns on human amniotic fluid profile. *Sci Rep*. 2018;8(1):4076. doi:10.1038/s41598-018-22230-y
83. Stocchero M, Locci E, d'Aloja E, Nioi M, Baraldi E, Giordano G. PLS2 in Metabolomics. *Metabolites*. 2019;9(3):51. doi:10.3390/metabo9030051

84. Sut S, Baldan V, Faggian M, Peron G, Dall`Acqua S. Nutraceuticals, A New Challenge for Medicinal Chemistry. *CMC*. 2016;23(28):3198-3223. doi:10.2174/0929867323666160615104837
85. Fernández-Peralbo MA, Luque de Castro MD. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC Trends in Analytical Chemistry*. 2012;41:75-85. doi:10.1016/j.trac.2012.08.011
86. Saude EJ, Sykes BD. Urine stability for metabolomic studies: effects of preparation and storage. *Metabolomics*. 2007;3(1):19-27. doi:10.1007/s11306-006-0042-2
87. Smith L, Villaret-Cazadamont J, Claus SP, et al. Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites*. 2020;10(3):104. doi:10.3390/metabo10030104
88. Savorani F, Rasmussen MA, Mikkelsen MS, Engelsen SB. A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. *Food Research International*. 2013;54(1):1131-1145. doi:10.1016/j.foodres.2012.12.025
89. Chen Y, Xu J, Zhang R, Abliz Z. Methods used to increase the comprehensive coverage of urinary and plasma metabolomes by MS. *Bioanalysis*. 2016;8(9):981-997. doi:10.4155/bio-2015-0010
90. Álvarez-Sánchez B, Priego-Capote F, Castro MDL de. Metabolomics analysis II. Preparation of biological samples prior to detection. *TrAC Trends in Analytical Chemistry*. 2010;29(2):120-127. doi:10.1016/j.trac.2009.12.004
91. Brennan L. NMR-based metabolomics: From sample preparation to applications in nutrition research. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 2014;83:42-49. doi:10.1016/j.pnmrs.2014.09.001
92. Gowda GAN, Djukovic D. Overview of Mass Spectrometry-Based Metabolomics: Opportunities and Challenges. In: Raftery D, ed. *Mass Spectrometry in Metabolomics*. Vol 1198. Methods in Molecular Biology. Springer New York; 2014:3-12. doi:10.1007/978-1-4939-1258-2_1
93. Emwas A-H, Roy R, McKay RT, et al. NMR Spectroscopy for Metabolomics Research. *Metabolites*. 2019;9(7):123. doi:10.3390/metabo9070123

94. Emwas A-HM, Salek RM, Griffin JL, Merzaban J. NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. *Metabolomics*. 2013;9(5):1048-1072. doi:10.1007/s11306-013-0524-y
95. Takis PG, Ghini V, Tenori L, Turano P, Luchinat C. Uniqueness of the NMR approach to metabolomics. *TrAC Trends in Analytical Chemistry*. 2019;120:115300. doi:10.1016/j.trac.2018.10.036
96. Yuliana ND, Jahangir M, Verpoorte R, Choi YH. Metabolomics for the rapid dereplication of bioactive compounds from natural sources. *Phytochem Rev*. 2013;12(2):293-304. doi:10.1007/s11101-013-9297-1
97. Pitt JJ. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin Biochem Rev*. 2009;30(1):19-34.
98. Banerjee S, Mazumdar S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry*. 2012;2012:1-40. doi:10.1155/2012/282574
99. Liigand P, Kaupmees K, Haav K, et al. Think Negative: Finding the Best Electrospray Ionization/MS Mode for Your Analyte. *Anal Chem*. 2017;89(11):5665-5668. doi:10.1021/acs.analchem.7b00096
100. Lei Z, Huhman DV, Sumner LW. Mass Spectrometry Strategies in Metabolomics. *Journal of Biological Chemistry*. 2011;286(29):25435-25442. doi:10.1074/jbc.R111.238691
101. Sadygov RG. High-Resolution Mass Spectrometry for In Vivo Proteome Dynamics using Heavy Water Metabolic Labeling. *IJMS*. 2020;21(21):7821. doi:10.3390/ijms21217821
102. Tugizimana F, Steenkamp PA, Piater LA, Dubery IA. Mass spectrometry in untargeted liquid chromatography/mass spectrometry metabolomics: Electrospray ionisation parameters and global coverage of the metabolome. *Rapid Commun Mass Spectrom*. 2018;32(2):121-132. doi:10.1002/rcm.8010
103. Want EJ, Cravatt BF, Siuzdak G. The Expanding Role of Mass Spectrometry in Metabolite Profiling and Characterization. *ChemBioChem*. 2005;6(11):1941-1951. doi:10.1002/cbic.200500151

104. Wolfender J-L, Marti G, Thomas A, Bertrand S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *Journal of Chromatography A*. 2015;1382:136-164. doi:10.1016/j.chroma.2014.10.091

105. Ren J-L, Zhang A-H, Kong L, Wang X-J. Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Adv*. 2018;8(40):22335-22350. doi:10.1039/C8RA01574K

106. Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: beyond biomarkers and towards mechanisms. *Nat Rev Mol Cell Biol*. 2016;17(7):451-459. doi:10.1038/nrm.2016.25

107. Periat A, Guillarme D, Veuthey J-L, et al. Optimized selection of liquid chromatography conditions for wide range analysis of natural compounds. *Journal of Chromatography A*. 2017;1504:91-104. doi:10.1016/j.chroma.2017.05.024

108. Fujito Y, Hayakawa Y, Izumi Y, Bamba T. Importance of optimizing chromatographic conditions and mass spectrometric parameters for supercritical fluid chromatography/mass spectrometry. *Journal of Chromatography A*. 2017;1508:138-147. doi:10.1016/j.chroma.2017.05.071

109. Zeki ÖC, Eylem CC, Reçber T, Kır S, Nemutlu E. Integration of GC–MS and LC–MS for untargeted metabolomics profiling. *Journal of Pharmaceutical and Biomedical Analysis*. 2020;190:113509. doi:10.1016/j.jpba.2020.113509

110. Seger C. Usage and limitations of liquid chromatography-tandem mass spectrometry (LC–MS/MS) in clinical routine laboratories. *Wien Med Wochenschr*. 2012;162(21-22):499-504. doi:10.1007/s10354-012-0147-3

111. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. Untargeted Metabolomics Strategies—Challenges and Emerging Directions. *J Am Soc Mass Spectrom*. 2016;27(12):1897-1905. doi:10.1007/s13361-016-1469-y

112. Jorge TF, Mata AT, António C. Mass spectrometry as a quantitative tool in plant metabolomics. *Phil Trans R Soc A*. 2016;374(2079):20150370. doi:10.1098/rsta.2015.0370

113. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics: MASS SPECTROMETRY-BASED METABOLOMICS. *Mass Spectrom Rev*. 2007;26(1):51-78. doi:10.1002/mas.20108

114. Koek MM, Jellema RH, van der Greef J, Tas AC, Hankemeier T. Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics*. 2011;7(3):307-328. doi:10.1007/s11306-010-0254-3

115. Vettukattil R. Preprocessing of Raw Metabonomic Data. In: Bjerrum JT, ed. *Metabonomics*. Vol 1277. Methods in Molecular Biology. Springer New York; 2015:123-136. doi:10.1007/978-1-4939-2377-9_10

116. Alsaleh M, Barbera TA, Andrews RH, et al. Mass Spectrometry: A Guide for the Clinician. *Journal of Clinical and Experimental Hepatology*. 2019;9(5):597-606. doi:10.1016/j.jceh.2019.04.053

117. Smolinska A, Blanchet L, Buydens LMC, Wijmenga SS. NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. *Analytica Chimica Acta*. 2012;750:82-97. doi:10.1016/j.aca.2012.05.049

118. Ren S, Hinzman AA, Kang EL, Szczesniak RD, Lu LJ. Computational and statistical analysis of metabolomics data. *Metabolomics*. 2015;11(6):1492-1513. doi:10.1007/s11306-015-0823-6

119. Martin M, Legat B, Leenders J, et al. PepsNMR for 1H NMR metabolomic data pre-processing. *Analytica Chimica Acta*. 2018;1019:1-13. doi:10.1016/j.aca.2018.02.067

120. van den Berg RA, Hoefsloot HC, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*. 2006;7(1):142. doi:10.1186/1471-2164-7-142

121. Tian H, Li B, Shui G. Untargeted LC–MS Data Preprocessing in Metabolomics. *J Anal Test*. 2017;1(3):187-192. doi:10.1007/s41664-017-0030-8

122. Mastrangelo A, Ferrarini A, Rey-Stolle F, García A, Barbas C. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Analytica Chimica Acta*. 2015;900:21-35. doi:10.1016/j.aca.2015.10.001

123. Saccenti E, Hoefsloot HCJ, Smilde AK, Westerhuis JA, Hendriks MMWB. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics*. 2014;10(3):361-374. doi:10.1007/s11306-013-0598-6

124. Chang H-Y, Colby SM, Du X, et al. A Practical Guide to Metabolomics Software Development. *Anal Chem*. 2021;93(4):1912-1923. doi:10.1021/acs.analchem.0c03581

125. Burns DC, Reynolds WF. Minimizing the risk of deducing wrong natural product structures from NMR data. *Magn Reson Chem*. 2021;59(5):500-533. doi:10.1002/mrc.4933

126. Bingol K. Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-Throughput*. 2018;7(2):9. doi:10.3390/ht7020009

127. Chaleckis R, Meister I, Zhang P, Wheelock CE. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Current Opinion in Biotechnology*. 2019;55:44-50. doi:10.1016/j.copbio.2018.07.010

128. Nordström A, Lewensohn R. Metabolomics: Moving to the Clinic. *J Neuroimmune Pharmacol*. 2010;5(1):4-17. doi:10.1007/s11481-009-9156-4

129. Bardanzellu F, Fanos V. How could metabolomics change pediatric health? *Ital J Pediatr*. 2020;46(1):37. doi:10.1186/s13052-020-0807-7

130. Mamas M, Dunn WB, Neyses L, Goodacre R. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Arch Toxicol*. 2011;85(1):5-17. doi:10.1007/s00204-010-0609-6

131. Greef J van der, Stroobant P, Heijden R van der. The role of analytical sciences in medical systems biology. *Current Opinion in Chemical Biology*. 2004;8(5):559-565. doi:10.1016/j.cbpa.2004.08.013

132. Tolstikov V, Moser AJ, Sarangarajan R, Narain NR, Kiebish MA. Current Status of Metabolomic Biomarker Discovery: Impact of Study Design and Demographic Characteristics. Metabolites. 2020;10(6):224. Published 2020 May 29. doi:10.3390/metabo10060224

133. Zhang A, Sun H, Yan G, Wang P, Wang X. Metabolomics for Biomarker Discovery: Moving to the Clinic. *BioMed Research International*. 2015;2015:1-6. doi:10.1155/2015/354671

134. Lin S-H, Fan J, Zhu J, et al. Exploring plasma metabolomic changes in sepsis: a clinical matching study based on gas chromatography–mass spectrometry. *Ann Transl Med*. 2020;8(23):1568-1568. doi:10.21037/atm-20-3562

135. Wang J, Sun Y, Teng S, Li K. Prediction of sepsis mortality using metabolite biomarkers in the blood: a meta-analysis of death-related pathways and prospective validation. *BMC Med*. 2020;18(1):83. doi:10.1186/s12916-020-01546-5

136. Ferrario M, Cambiaghi A, Brunelli L, et al. Mortality prediction in patients with severe septic shock: a pilot study using a target metabolomics approach. *Sci Rep*. 2016;6(1):20391. doi:10.1038/srep20391

137. Liu Z, Triba MN, Amathieu R, et al. Nuclear magnetic resonance-based serum metabolomic analysis reveals different disease evolution profiles between septic shock survivors and non-survivors. *Crit Care*. 2019;23(1):169. doi:10.1186/s13054-019-2456-z

138. Liu Z, Yin P, Amathieu R, Savarin P, Xu G. Application of LC-MS-based metabolomics method in differentiating septic survivors from non-survivors. *Anal Bioanal Chem*. 2016;408(27):7641-7649. doi:10.1007/s00216-016-9845-9

139. Neugebauer S, Giamarellos-Bourboulis EJ, Pelekanou A, et al. Metabolite Profiles in Sepsis: Developing Prognostic Tools Based on the Type of Infection*. *Critical Care Medicine*. 2016;44(9):1649-1662. doi:10.1097/CCM.0000000000001740

140. Su L, Huang Y, Zhu Y, et al. Discrimination of sepsis stage metabolic profiles with an LC/MS-MS-based metabolomics approach. *BMJ Open Resp Res*. 2014;1(1):e000056. doi:10.1136/bmjresp-2014-000056

141. Seymour CW, Yende S, Scott MJ, et al. Metabolomics in pneumonia and sepsis: an analysis of the GenIMS cohort study. *Intensive Care Med*. 2013;39(8):1423-1434. doi:10.1007/s00134-013-2935-7

142. Langley RJ, Tsalik EL, Velkinburgh JC van, et al. An Integrated Clinico-Metabolomic Model Improves Prediction of Death in Sepsis. *Sci Transl Med*. 2013;5(195). doi:10.1126/scitranslmed.3005893

143. Kauppi AM, Edin A, Ziegler I, et al. Metabolites in Blood for Prediction of Bacteremic Sepsis in the Emergency Room. Yende S, ed. *PLoS ONE*. 2016;11(1):e0147670. doi:10.1371/journal.pone.0147670

144. Elmassry MM, Mudaliar NS, Colmer-Hamood JA, et al. New markers for sepsis caused by Pseudomonas aeruginosa during burn infection. *Metabolomics*. 2020;16(3):40. doi:10.1007/s11306-020-01658-2

145. Mickiewicz B, Vogel HJ, Wong HR, Winston BW. Metabolomics as a Novel Approach for Early Diagnosis of Pediatric Septic Shock and Its Mortality. *Am J Respir Crit Care Med*. 2013;187(9):967-976. doi:10.1164/rccm.201209-1726OC

146. Mickiewicz B, Thompson GC, Blackwood J, et al. Development of metabolic and inflammatory mediator biomarker phenotyping for early diagnosis and triage of pediatric sepsis. *Crit Care*. 2015;19(1):320. doi:10.1186/s13054-015-1026-2

147. Dessì A, Liori B, Caboni P, et al. Monitoring neonatal fungal infection with metabolomics. *The Journal of Maternal-Fetal & Neonatal Medicine*. 2014;27(sup2):34-38. doi:10.3109/14767058.2014.954787

148. Sarafidis K, Chatziioannou AC, Thomaidou A, et al. Urine metabolomics in neonates with late-onset sepsis in a case-control study. *Sci Rep*. 2017;7(1):45506. doi:10.1038/srep45506

149. Pandey S, Adnan Siddiqui Mohd, Azim A, Trigun SK, Sinha N. Serum metabolic profiles of septic shock patients based upon co-morbidities and other underlying conditions. *Mol Omics*. 2021;17(2):260-276. doi:10.1039/D0MO00177E

150. Zhang P, Wang Z, Qiu H, Zhou W, Wang M, Cheng G. Machine learning applied to serum and cerebrospinal fluid metabolomes revealed altered arginine metabolism in neonatal sepsis with meningoencephalitis. *Computational and Structural Biotechnology Journal*. 2021;19:3284-3292. doi:10.1016/j.csbj.2021.05.024

151. Report on the Expert Meeting on Neonatal and Paediatric Sepsis. EMA London 8 June 2010. Available online: https://www.ema.europa.eu/en/events/expert-meeting-neonatal-paediatric-sepsis

152. Santamaria F, Montella S, Stocchero M, et al. Effects of pidotimod and bifidobacteria mixture on clinical symptoms and urinary metabolomic profile of children with recurrent respiratory infections: a randomized placebo-controlled trial. *Pulmonary Pharmacology & Therapeutics*. 2019;58:101818. doi:10.1016/j.pupt.2019.101818

153. Stocchero M. Relevant and irrelevant predictors in PLS2. *Journal of Chemometrics*. 2020;34(8). doi:10.1002/cem.3237

154. Viant MR, Kurland IJ, Jones MR, Dunn WB. How close are we to complete annotation of metabolomes? *Current Opinion in Chemical Biology*. 2017;36:64-69. doi:10.1016/j.cbpa.2017.01.001

155. Bohár Z, Toldi J, Fülöp F, Vécsei L. Changing the Face of Kynurenines and Neurotoxicity: Therapeutic Considerations. *IJMS*. 2015;16(12):9772-9793. doi:10.3390/ijms16059772

156. Wlodarska M, Luo C, Kolde R, et al. Indoleacrylic Acid Produced by Commensal Peptostreptococcus Species Suppresses Inflammation. *Cell Host & Microbe*. 2017;22(1):25-37.e6. doi:10.1016/j.chom.2017.06.007

157. Hendrikx T, Schnabl B. Indoles: metabolites produced by intestinal bacteria capable of controlling liver disease manifestation. *J Intern Med*. 2019;286(1):32-40. doi:10.1111/joim.12892

158. Stewart CJ, Embleton ND, Marrs ECL, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome*. 2017;5(1):75. doi:10.1186/s40168-017-0295-1

159. Collado MC, Cernada M, Neu J, Pérez-Martínez G, Gormaz M, Vento M. Factors influencing gastrointestinal tract and microbiota immune interaction in preterm infants. *Pediatr Res*. 2015;77(6):726-731. doi:10.1038/pr.2015.54

160. Madan JC, Salari RC, Saxena D, et al. Gut microbial colonisation in premature neonates predicts neonatal sepsis. *Arch Dis Child Fetal Neonatal Ed*. 2012;97(6):F456-F462. doi:10.1136/fetalneonatal-2011-301373

161. Lyons J, Rauh-Pfeiffer A, Ming-Yu Y, et al. Cysteine metabolism and whole blood glutathione synthesis in septic pediatric patients: *Critical Care Medicine*. 2001;29(4):870-877. doi:10.1097/00003246-200104000-00036

162. Ambroggio L, Florin TA, Shah SS, et al. Emerging Biomarkers of Illness Severity: Urinary Metabolites Associated with Sepsis and Necrotizing Methicillin-Resistant *Staphylococcus aureus* Pneumonia. *Pharmacotherapy*. 2017;37(9):1033-1042. doi:10.1002/phar.1973

163. Bhandari V, Wang C, Rinder C, Rinder H. Hematologic Profile of Sepsis in Neonates: Neutrophil CD64 as a Diagnostic Marker. *PEDIATRICS*. 2008;121(1):129-134. doi:10.1542/peds.2007-1308

164. Benitz WE. Adjunct Laboratory Tests in the Diagnosis of Early-Onset Neonatal Sepsis. *Clinics in Perinatology*. 2010;37(2):421-438. doi:10.1016/j.clp.2009.12.001

165. Al-Zahrani AK, Ghonaim MM, Hussein YM, Eed EM, Khalifa AS, Dorgham LS. Evaluation of recent methods versus conventional methods for diagnosis of early-onset neonatal sepsis. *J Infect Dev Ctries*. 2015;9(04):388-393. doi:10.3855/jidc.5950