A six molecule nenonatal serum metabolite biosignature has high neonatal sepsis predictability and resolves upon treatment completion.

Riya Ahmed[1], Anil Behera[1], Adyasha Sarangi[1], Pradeep Debata[2], Rajni Gaind[3], GP Kaushal[4], Renu Gur[5], Sushil Shrivastava[6], Kirti Nirmal[7], Ravinder Kaur[8], Sushma Nangia[9], Vivek Kumar[10], M Jeeva Sankar[10,*], Ranjan Kumar Nanda[1,*]

[1]Translational Health Group, International Center for Genetic Engineering and Biotechnology, New Delhi, India, 110067.

[2]Department of Paediatrics, Vardhman Mahavir Medical College and Safdarjung Hospital India, New Delhi, India, 110029.

[3]Department of Microbiology, Vardhman Mahavir Medical College and Safdarjung Hospital India, New Delhi, India, 110029.

[4]Department of Paediatrics, Dr. Baba Saheb Ambedkar Hospital, New Delhi, India, 110085.

[5]Department of Microbiology, Dr. Baba Saheb Ambedkar Hospital, New Delhi, India, 110085.

[6]Department of Paediatrics, Guru Teg Bahadur Hospital, New Delhi, India, 110095.

[7]Department of Microbiology, Guru Teg Bahadur Hospital, New Delhi, India, 110095.

[8]Department of Microbiology, Lady Hardinge Medical College and Associated Hospitals, New Delhi, India, 110001.

[9]Department of Neonatology, Lady Hardinge Medical College and Associated Hospitals, New Delhi, India, 110001.

[10]Department of Paediatrics, All India Institute of Medical Sciences, New Delhi, India, 110029.

***Corresponding Authors' contact:**
Ranjan Kumar Nanda (PhD)
Translational Health Group
International Centre for Genetic Engineering and Biotechnology (ICGEB)
New Delhi Component, Aruna Asaf Ali Marg
INDIA-110067
Telephone: +91-11-26741358 -426
Fax: +91-11-26742316
E-mail: ranjan@icgeb.res.in
and/or
M Jeeva Sankar (MD)
Department of Paediatrics
All India Institute of Medical Science (AIIMS)
New Delhi, Sri Aurobindo Marg, Ansari Nagar
INDIA-110029
E-mail: jeevasankar@gmail.com

ABSTRACT

Background: Sepsis, a life-threatening disorder with multi-organ dysfunction, is a leading cause of neonatal mortality. Current microbiology-based sepsis diagnosis is time-consuming, and identification of deregulated host serum metabolite signatures might be useful to develop early screening tools and host-directed therapeutics.

Methods: In this multi-institutional study, 500 neonates (41.2% female) were classified to culture-positive (CP) or negative sepsis (CN) cases and controls (no sepsis: NS, healthy control: HC) based on their microbial culture and mass spectrometry test results. The neonates were randomly grouped into two discovery sets (I:n=71; II:n=269), a validation set (n=60), and a longitudinally followed-up population (n=100). Serum samples of these neonates were processed and profiled using gas chromatography coupled to either quadrupole or time-of-flight mass spectrometry (GC-MS/-TOF-MS). Deregulated (log_2case/control $\geq \pm 0.58$, $p<0.05$) serum metabolites in sepsis cases were identified from the discovery sets and their predictive accuracy in the validation set was calculated using area under the receiving operator characteristic curve (AUC of ROC). The abundance of these deregulated metabolites was monitored in the longitudinally followed-up neonates (CP:n=29, CN:n=35, and NS:n=36) completing therapeutic intervention.

Results: Most of the CP cases were *Klebsiella pneumoniae* (28.6%) or *Acinetobacter baumannii* (20.6%) positive. Gestational age (CP: 30.9±1.9 weeks, CN: 30.9±1.8 weeks, HC: 32.3±1.3 weeks, NS: 31.6±1.5 weeks) and birthweight (CP: 1.4±0.3 kg, CN: 1.4±0.4 kg, HC: 1.7±0.3 kg, NS: 1.6±0.3 kg) were lower in sepsis neonates compared to controls. Out of 57 identified serum metabolites, a set of six (1,5-Anhydro-D-sorbitol-Lactic-acid-Malic-acid-Myo-inositol-Phenylalanine-Lysine) were identified as sepsis biosignature. The AUC of ROC of the biosignature to predict CP or CN from HC was 0.97 and from NS was 0.84 and 0.64, respectively. Myo-inositol, malic acid, and 1,5-anhydro-D-sorbitol revert to the HC levels in neonates completing therapeutic intervention.

Conclusions: A serum metabolite signature showed a >97% predictive accuracy for sepsis and could be further explored for its diagnostic and host-directed therapeutic potential.

Keywords: Neonatal sepsis, metabolomics, biomarker, clinical study, GC-MS

INTRODUCTION

In 2017, ~62 million new sepsis cases with 11 million deaths were reported globally, and every year 1-2 million new cases are reported with an 11-19% mortality rate (*1*, *2*). Sepsis presents as a life-threatening multi-organ dysfunction caused by a dysregulated host response to infection caused by pathogens like *Acinetobacter baumannii*, *Escherichia coli,* and *Klebsiella spp* (*3*, *4*). Culture-positive sepsis has been linked to a higher risk of death and requires longer hospital stays and mechanical ventilation compared to those with culture-negative sepsis (*5*). Due to their underdeveloped immune system and gut microbiota, neonatal sepsis may have different presentations and earlier diagnosis will save their lives by appropriate intervention. (*6*).

Sepsis-induced aggressive immune response may influence metabolic pathways in affected cells and tissues and reflect in the systemic circulation. Currently, serum procalcitonin and C-reactive protein (CRP) are used biomarkers for clinical diagnosis of sepsis but suffer from low specificities (*7*, *8*). Other biomarkers like CD64, CD11b, IP10, IL-12, and sTREM-1 have a higher predictive ability in adult sepsis (*9*). In neonates, a combination of ratio of immature to total neutrophil ratio (I/T ratio), CRP levels, total blood count, and total leucocyte count-erythrocyte sedimentation rate (TLC ESR) are used for sepsis screening. Blood lactate and levels were reported to have a direct correlation with disease severity and mortality in adult sepsis shock patients however, their combined predictive ability have not been reported (*10*, *11*). Mass spectrometry-based global metabolic profiling of biofluids of sepsis patients and controls might help identify key deregulated metabolites to better understand the disease pathophysiology and diagnostic biosignatures.

In this study, serum metabolites of neonates with sepsis and controls (suspected but later labelled as having no sepsis as well as healthy controls) were profiled using mass spectrometry to identify putative markers and they were monitored in longitudinally followed-up neonatal sepsis patients completing therapeutic interventions. Identification and validation of a sepsis serum molecular signature will help us develop novel diagnostic tools to classify the patients timely and a better understanding of the perturbed pathways in sepsis will be useful in the development of host-directed therapeutics.

MATERIALS AND METHODS

Subject recruitment and classification: The study was approved by the Institutional Ethics Committees of All India Institute of Medical Sciences (IEC-1074/06.11.20, RP-28/2020), Vardhman Mahavir Medical College (VMMC) and Safdarjung Hospital, New Delhi (IEC/VMMC/SJH/Project/2021-05/CC-157, dated 24.07.2021), Dr. Baba Saheb Ambedkar Medical College and Hospital (5(32)/2020/BSAH/DNB/PF, dated 31.08.2020), Lady Hardinge Medical College (LHMC/IEC/2022/03/30), University College of Medical Sciences & Guru Teg Bahadur Hospital (IECHC-2022-51-R1, dated 07.12.2022) and International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi (ICGEB/IEC/2021/28, version 2). Preterm neonates born between 28 and 34 weeks of gestation and admitted to the neonatal intensive care unit (NICU) of the clinical sites from July 2022 to December 2023 were enrolled

after obtaining written consent from the legal guardians. Neonates born to HIV +ve or hepatitis +ve mothers or postnatal age of > 28 days were excluded (Figure S1, Table S1). Neonates born without any complications or symptoms of sepsis were recruited as healthy controls (HC) from each clinical site.

The clinical team suspected sepsis in the presence of either maternal or perinatal risk factors or based on clinical signs supported by dedicated research staff to the enrolled neonates on a 24×7 basis (Table S1). The research nurses performed the diagnostic work-up for sepsis, including a sepsis screen, at the time of sepsis suspicion and before initiating antibiotics. Whole blood sample (1mL) was collected from the neonates for culture in automated culture bottles (BD BACTEC or BACT/ALERT) and incubated at 37° C before being transported to the microbiology laboratory within 12 to 24 hours of collection.

Based on the laboratory test results and the Centre for Disease Control and Prevention's (CDC) National Healthcare Safety Network (NHSN) infection definitions, the neonates were grouped as having culture-positive (CP) sepsis, culture-negative (CN) sepsis, no sepsis (NS), or healthy controls (HC)(*12*) (Table S2). Primarily, neonates with a positive screen whose whole blood culture grew a true pathogen were classified as CP and if the culture was sterile or grew commensals they were grouped as CN. The sepsis screen included: white blood cell count < 4.0×10^9 cells/L; absolute neutrophil count < 1.5×10^9 cells/L; I/T ratio > 0.2; CRP > 6mg/L; TLC ESR > 15mm in 1st hour; the screen was considered positive in the presence of at least two abnormal findings (*13*). Neonates suspected of sepsis but not meeting the criteria for either CP or CN sepsis were classified as NS (Tables S1 and S2). The decision to start antibiotics was up to the clinical team's discretion once the samples for the sepsis workup were drawn following their unit policy.

The enrolled neonates were randomly distributed into the discovery set-I (n=71, ~59.2% female, CP/CN/NS/HC: 16/24/17/14), discovery set-II (n=269, ~43.4% female, CP/CN/NS/HC: 63/69/68/69), and validation set (n=60, ~40% female, CP/CN/NS/HC: 15 each) using an online randomizer (randomizer.org). A set of neonates (n=100, CP/CN/NS: 29/35/36) independent from the earlier sets were longitudinally followed up till recovery to monitor the abundance of the deregulated metabolites.

Serum Sample Processing: A part of the whole blood (0.5 ml) was transferred to serum collection tubes and kept still for 10-20 minutes. After centrifuging at 1000 × g for 10 minutes at 4℃, the supernatant was transferred into fresh tubes before storing at -80℃. A quality control (QC) sample for each set was prepared separately by pooling equal volumes of serum samples. All samples from the participating subjects were randomized using the online tool and run in batches of 20 with at least 2 QC samples per batch.

Serum sample processing and metabolite extraction: A set of randomized stored serum samples (n=20) with 2 QCs were thawed at 4℃ and to an aliquot (40 µl) of it, chilled methanol (80%, 320 µl) was added with a spike in standard (ribitol: 0.5 µg/µl, 1 µl). After incubating it in ice for 30 minutes, it was centrifuged at 4℃ for 10 minutes at 15,000 × g. The supernatant was transferred to a fresh methanol-treated centrifuge tube. A blank tube was used as extraction

blank with spiked in ribitol. After drying the samples in a Speedvac (Labconco, USA), they were derivatized using N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA). After adding MOX-HCl solution (20 mg/mL; 40 µl) to each sample tube, they were incubated at 60℃ for 2 hours at 900 × g. Then MSTFA (70 µl) was added to each sample tube and incubated at 60℃ for 30 minutes at 900 × g. The derivatized samples were centrifuged at 16,000 × g for 10 minutes at room temperature and the supernatant (~90 µl) was transferred to fresh GC vials with 250 µl glass inserts for analysis using different GC-MS platforms.

GC-MS Data acquisition: The discovery set-I samples were analyzed using a GC coupled to a quadrupole-MS (the Agilent 5975c TAD series). Using helium gas at a constant flow rate of 1 ml/minute, the derivatized sample (1 µl) was injected in splitless mode to an RTX-5 column (30 m by 0.25 mm by 0.25 µm; Restek, USA). The oven temperature gradient was set from 50℃ to 280℃ with a ramp of 8.5℃/minute from 50 to 200℃ and 6℃/minute from 200 to 280℃ with a hold of 5 minutes for metabolite separation. The temperature of the MSD transfer line was set at 260℃ with a holding time of 1.37 minutes. The detector voltage was set at 1388 V and a mass range of 50 to 600 m/z was selected. The MS ion source and quadrupole temperatures were set at 230℃ and 150℃, respectively. A solvent delay of 10 minutes was set, and the total GC-MS run time was 36.98 minutes.

For discovery set-II, validation, and follow-up sets, the derivatized test samples were analyzed using a GC coupled to a time-of-flight(TOF)-MS instrument (Pegasus 4D; Leco, USA). Using helium gas at a constant flow rate of 1 ml/minute, the derivatized sample (1 µl) was injected in splitless mode to an RTX-5 column (30 m by 0.25 mm by 0.25 µm; Restek, USA). The oven temperature gradient was set from 50℃ to 280℃ with a ramp of 8.5℃/minute from 50 to 200℃ and 6℃/minute from 200 to 280℃ with a hold time of 5 minutes for metabolite separation. The detector voltage was set at 1400 V and a mass range of 30 to 600 m/z was used for detection. The temperature of the MSD transfer line was set at 250℃, and the flow rate was set at 1 ml/minute. The MS ion source and acquisition rate were set at 220℃ and 30 spectra/second, respectively. The total GC-MS run time was 36.98 minutes with a 10 minute solvent delay.

Data Pre-processing: The .d files from the discovery set-I were then imported to the ChromaTOF® (LECO, USA) software for further analysis. The raw data files obtained from each data set were separately analysed in the ChromaTOF® (LECO, USA) software using the 'Statistical Compare' feature. Data matrices consisting of identified metabolites as variables and the peak areas identified in the samples of each dataset were created. NIST 11 (National Institute of Standards and Technology, 243893 spectra; 30932 replicates) Mass Spectral database and LECO/Fiehn Metabolomics Library (ChromaTOF, USA) were used for the tentative identification of the molecular features. The minimum similarity match was set to 600 for annotation. The identification of significantly deregulated features was further confirmed by running commercial standards using the same GC-MS parameters.

Multivariate and univariate statistical analysis: Clinical parameters with $p<.05$ at 95% confidence interval between study groups were selected as significant. MetaboAnalyst 4.0, GraphPad Prism 9.0.0 (GraphPad Software, Boston, Massachusetts, USA), and RStudio were

used to identify deregulated metabolites (*14*). Principal component analysis (PCA) and partial least square discriminate analysis (PLS-DA) models were generated separately for each dataset using MetaboAnalyst. A *t*-test with Welch's correction was performed to compare the differences in the abundance of all the metabolites between groups, using normalized data. A paired *t*-test was performed for the longitudinally follow-up data to monitor the abundance of the deregulated metabolites. All identified metabolites were selected for Metabolite Gene Set Analysis (MSEA) using MetaboAnalyst to identify the perturbed molecular pathways. The area under the curve of Receiver Operating Characteristic (AUC of ROC) was calculated using the "combiroc" package in R. For visualization of the resulting data, different softwares – GraphPad Prism 9, Microsoft PowerPoint, and R packages "ggplot2" and "scatter3d" – were employed. An overview of the workflow is represented in Figure 1. Metabolites with a fold change value >1.5 and *t*-test p-value <0.05 were considered as significantly deregulated metabolites.

RESULTS

Clinical details of the study subjects: A set of 500 neonates (~41.2% female) belonging to sepsis (CP: n=124, CN: n=144) and controls (NS: n=134, HC: n=98) were included in this study. *Acinetobacter baumannii* was detected in the whole blood of ~37.5% of the CP cases, and coagulase-negative *Staphylococcus* (CoNS) was identified in ~19% of the CP cases in the discovery set-I. In the discovery set-II, *Acinetobacter baumannii* was present in 20.6% and *Klebsiella pneumonia* was present in 28.6% of the CP cases. In the validation set, *Acinetobacter baumannii* was found in 33.3% of the CP cases. The detailed clinical parameters of the neonates are presented in Additional File A1. In the discovery set-I, the gestational age of neonates of CP (30.9±1.7 weeks) and CN (30.8±2.1 weeks) sepsis groups was significantly lower than that of the HC group (32.9±1.5 weeks). In discovery set-II, the gestational age was significantly lower for sepsis groups (CP/CN: 30.9±1.9 weeks/30.9±1.2 weeks) compared to the controls (HC/NS: 32.3±1.3 weeks/31.6±1.5 weeks). The gestational age was also lower in the sepsis groups (CP/CN: 30.9±2 weeks/30.5±1.9 weeks) compared to the control groups (HC/NS: 32.6±1.5 weeks/32.3±1.5 weeks). Similarly, the birth weight of the neonates of CP and CN groups of the discovery set-I (CP/CN: 1.3±0.5 kg/1.4±0.3 kg) and discovery set-II (CP/CN: 1.4±0.3 kg/1.5±0.4 kg) was found to be lower than the healthy controls (2±0.4 kg). In the validation set, birth weight was lower in CN (1.3±0.3 kg) compared to the HC group (1.7±0.3 kg). In the discovery set-I, the number of cesarean deliveries was lower in the CN group (23.8%) as compared to HC (43.8%). In discovery set-II, the number of cesarean deliveries was lower in CP (46.7%) than in the other groups. Mortality in sepsis groups (CP/CN: 77.3%/65%) of the validation set was significantly higher than the NS group (27.9%). The absolute neutrophil count was found to be significantly lower in the CP ($3.8\times10^3\pm2.5\times10^3$) and CN ($3.7\times10^3\pm2.5\times10^3$) groups compared to the NS ($6.9\times10^3\pm3.4\times10^3$) group of discovery set-I. This difference in absolute neutrophil count was limited in the discovery set-II and validation set. The group-specific clinical details are presented in Table 1 and S3.

Serum metabolic phenotyping of sepsis patients and controls using mass spectrometry: Serum metabolite profiling of neonates identified a set of 882 metabolic features, and PCA analysis showed a closed cluster of all the QC samples indicating minimum method-associated

variations (Figure 2A). A set of 52 metabolites fulfilled the criteria and a few samples from each group (CP/CN/NS/HC: 1/2/2/1) were outliers and removed from further analysis (Figure 2A). The PLS-DA plots clustered the CP and CN groups away from the control (HC and NS) groups (Figure 2B, 2C, 2D, S2A).

A signature of 7 deregulated metabolites ($log_2FC>0.58$, $p<0.05$) (phenylalanine, diethyl phthalate, histidine, myo-inositol, serine, tyrosine, lysine) were identified in CP compared to the HC group (Figure 2B, S2B). The serum metabolic phenotype of CP and NS groups was found to be similar (Figure S2C). A set of 9 deregulated metabolites (histidine, myo-inositol, serine, L-tyrosine, L-lysine, 19-norethindrone, oleic acid, galactitol, maltose,) was identified in the CN group compared to the HC group (Figure 2C, S2B, S2C). One metabolite (1,5-anhydro-D-sorbitol) was found to be deregulated in CN compared to the NS group (Figure S1B). A signature of 5 metabolites (linoleic acid, 1,5-anhydro-D-sorbitol, 5α-androst-2-ene-17-one, threonine, 19-norethindrone) could differentiate between CP and CN groups (Figure 2D, S2D, S2E).

Serum metabolite profiling of discovery set-II: GC-TOF metabolite profiling of neonatal serum of discovery set-II identified 559 metabolic features, out of which 57 qualified the set criteria, and minimum method-associated variation was observed (Figure 2E). A few samples from each study group (CP:1, CN:1; NS:1, HC:1) were identified as outliers and were excluded from further analysis. PLS-DA model showed clear cluster of CP and CN study groups compared to the HC and NS groups (Figure 2F, 2G, 2H, S3A).

A set of 19 deregulated metabolites (1,5-anhydro-D-sorbitol, 2-hydroxybutyric acid, 3-aminoisobutyric acid, 3-hydroxybutyric acid, alanine, creatinine, cystein, lactic acid, gluconic acid, gluconic acid-e-lactone, indole-3-acetic acid, lysine, malic acid, myo-inositol, oxalic acid, phenylalanine, phosphoric acid, xylose, Unknown 1) could differentiate CP from the HC group (Figure 2F, S2B, S2E, S4). Three metabolites (lysine, xylohexulose, sorbitol) were deregulated in CP compared to NS (Figure S3C). Thirty-one significantly differential metabolites were identified between CN and HC groups (Figure 2G, S3B). One metabolite (indole-3-acetic acid) was found to be deregulated in CN compared to NS (Figure S3C). Eleven metabolites were deregulated between the CP and CN groups (Figure 2H, S3D, S3E, Table S4). MSEA of the identified deregulated metabolites revealed a perturbed inositol phosphate, ascorbate, and pyruvate metabolism in sepsis (CP/CN) groups compared to HC (Figure S4).

Predictive ability of the identified biosignatures on a test set (n=60): The AUCs of ROC of the 6 serum metabolite signature 1,5-Anhydro-D-sorbitol-Lactic-acid-Malic-acid-Myo-inositol-Phenylalanine-Lysine in CP were 0.97 and 0.84 against HC and NS, respectively (Figure 3A, S6A,). In CN, the AUC of ROC was 0.97 against HC and 0.64 against NS (Figure 3B, S6B). The metabolite signature of Sorbitol-Lysine-3-Hydroxybutyric-acid-Leucine-Aminomethane-Threonine-Indole-3-pyruvic-acid-Isoleucine-Xylo-5-hexosulose-Alanine could differentiate CP from CN with an AUC of 0.93 (Figure 3C).

Trend of serum biosignatures in longitudinally followed-up culture-positive/-negative sepsis patients: In the longitudinally followed-up CP and CN cases, the recovery time of CP

and CN cases was higher than NS (Figure 1). The levels of serum 1,5-anhydro-D-sorbitol ($p<0.0001$) and myo-inositol (CP: $p=0.002$, CN: $p=0.01$) significantly declined upon recovery. Serum phenylalanine and malic acid levels were decreased in recovered CN and NS subjects but not in CP sepsis patients. Serum malic acid levels had an increased trend in recovered CP cases, however, the difference was insignificant (Figure 3D, S7).

DISCUSSION

Neonatal sepsis is of mono or polymicrobial origin and leads to a complex host immune response with symptoms like anaphylaxis and euglycemic ketoacidosis (*15*). Neonatal sepsis mostly occurs in utero transplacentally or during delivery when the fetus comes in contact with the vaginal microenvironment. The hospital environment is known to harbor a myriad of microorganisms that contribute to sepsis cases. The presence of pathogens or their molecules contributes to alterations in host cell and tissue metabolism by altering their biochemical activities. These changes may be reflected in the metabolite phenotype of different biofluids like serum, urine, and cerebro-spinal fluid (CSF), which have been profiled using mass spectrometry (*16–18*). Identifying the deregulated metabolites specific to sepsis and their associated pathways can help us understand the pathology and develop appropriate interventions.

The gestational period of neonatal sepsis cases was lower compared to healthy controls, corroborating earlier reports (*19*). Infection rates have been reported to be inversely related to gestational age, with most sepsis neonates born before 32 weeks of gestation (*20*). Intrauterine infections and fetal inflammatory responses are known to cause preterm births which can explain the lower gestational age observed in our study (*21*). In the CP sepsis cases, *Klebsiella pneumonia* and *Acinetobacter baumannii* were the most common causative agents and they are known to be a primary source of infection in developing countries like India and Jordan, while Group B *Streptococcus* (GBS) is known to be more prevalent in developed countries like the US, UK, Ireland, and Portugal (*22–25*). Based on the results obtained in this study, we observed higher sensitivity and range in GC-TOF-MS (discovery set-II) compared to GC-MS (discovery set-I), because of which data acquisition of the validation set and follow-up samples was performed using GC-TOF-MS.

We report a biosignature of 1,5-anhydro-D-sorbitol, lactic acid, lysine, malic acid, myo-inositol, and phenylalanine. All these metabolite levels were elevated in CP and CN sepsis cases in comparison to the healthy controls. Phenylalanine metabolism is known to be upregulated in adult sepsis patients and positively associated with a higher risk of mortality (*11*, *26*, *27*). Higher serum and plasma lysine levels are observed in adult sepsis patients and the urine of neonates with sepsis (*28–30*). Lysine and phenylalanine are essential amino acids and their uptake in neonates is essential for proper growth and development. Their increased abundance in sepsis cases may be linked to intestinal dysbiosis that occurs in sepsis (*31*). Increased intestinal permeability also enhances the uptake of metabolites into the bloodstream. Impaired liver phenylalanine hydroxylase (PAH) enzyme, which catalyzes the conversion of phenylalanine to tyrosine, might also contribute to increased serum phenylalanine levels in sepsis. This could be due to PAH's dependency on its cofactor 5,6,7,8-tetrahydrobiopterin

(BH4), which is sensitive to oxidative stress and tends to be depleted in inflammatory conditions (*32*). Sepsis leads to muscle atrophy and the high serum lysine levels may be due to low cellular uptake since lysine-deficient apoptosis has been known to induce muscle protein degradation in muscle cells (*33*, *34*).

High urine and serum myo-inositol levels were reported in neonatal and geriatric sepsis patients, respectively (*35*). Plasma myo-inositol levels are known to be positively correlated with the acute physiology score (APS) in sepsis-induced acute lung injury (ALI) patients (*36*). Myo-inositol aids in the depolarization of the macrophage membrane potential which enhances their phagocytic ability (*37*). It is involved in the production of phosphatidylinositol (4,5)-bisphosphate (PIP2) and phosphatidylinositol (3,4,5)-trisphosphate (PIP3), which, in turn, are associated with the Akt pathway. The Akt pathway activates macrophages and various other immune cells and regulates apoptosis and oxidative stress in sepsis-induced tissue injuries (*38*, *39*). However, it is uncertain whether serum myo-inositol levels can directly reflect on the cellular activity of Akt pathway. Serum 1,5-Anhydro-D-sorbitol levels negatively correlate with the blood glucose levels of diabetic patients (*40*). This is because 1,5-anhydro-D-sorbitol, a 1-deoxy form of glucose, is known to be competitively inhibited by glucose in renal reabsorption (*41*). Uptake of 1,5-Anhydro-D-sorbitol has been previously reported to revert neonatal complications arising from diabetic mothers (*42*). 1,5-Anhydro-D-sorbitol is reported to block the production of pro-inflammatory cytokines like interleukin-6 (IL-6) and monocyte chemoattractant protein-1 (MCP-1) in murine macrophages treated with lipopolysaccharide (LPS) (*43*). The higher levels of 1,5-anhydro-D-sorbitol in the CP and CN sepsis groups could result from anti-inflammatory responses in host immune cells against sepsis. Increased glycolysis upon infection may also contribute to high serum 1,5-anhydro-D-sorbitol levels upon the increase in cellular glucose uptake (*44*).

Higher serum lactic acid levels are reported in sepsis patients and are directly associated with severity and mortality (*45*). This is due to the Warburg effect in inflamed tissues and immune cells where aerobic glycolysis is utilized to produce ATP to meet higher energy demands. The conversion of pyruvate to lactic acid is also enhanced to produce NADH which contributes towards glycolysis (*46*). Lactic acid has been reported to favour M2-like macrophage polarization (*47*). High plasma malic acid levels were reported in patients with sepsis-associated encephalopathy (*48*). Apart from its role as a TCA cycle intermediate in the mitochondria, malate gets transported into the cytosol where it can be converted to pyruvate which in turn contributes to increased lactate production during inflammation (*45*, *49*). The malate-pyruvate shuttle also aids in the production of NADPH which is essential for glycolysis (*50*). LPS exposure is reported to result in the accumulation of malate, fumarate, and succinate in macrophages (*51*).

It is important to highlight that the elevated levels of phenylalanine, myo-inositol, malic acid, and 1,5-anhydro-D-sorbitol in CP, CN, and NS sepsis neonates lowered upon treatment completion and recovery. The higher recovery time observed in CP and CN groups may be contributed by the onset of hyper-inflammmation and subsequent immunoparalysis which are difficult to resolve (*52*). This validates their association with sepsis onset and can be targeted for therapeutics development.

One of the limitations of this study is that the origin of the identified metabolites could not be traced back to either the host or the infectious agents. Identifying the enzymes catalyzing the deregulated pathways will be useful in understanding the mechanisms behind metabolite deregulation. Screening the mothers for their glucose levels and diabetes status could have helped us build a correlation between neonatal 1,5-anhydro-D-sorbitol levels and diabetic mothers.

In conclusion, a serum biosignature of myo-inositol, phenylalanine, malic acid, lysine lactic acid, and 1,5-anhydro-D-sorbitol could differentiate the CP and CN groups from HC and NS groups with high accuracy. The perturbed metabolic pathways in sepsis could be targeted for developing host-directed therapeutics.

ACKNOWLEDGEMENTS

We would like to thank the participants for their consent to collect blood for the project. We thank the hospital staff of the clinical sites for helping us with sample collection and preprocessing. We would like to acknowledge Dr. Meetu Salhan, Dr. Harsh Chellani, Dr. Pratima Anand, and Dr. Narendra Pal Singh for their constant support and involvement in this project. We would also like to thank the Department of Biotechnology (DBT), India, for providing financial support. We also thank the Translational Health group members for their help and support.

AUTHOR CONTRIBUTIONS

Conceptualization: R.N. Experiment design: R.N., R.A. Methodology: A.S., R.A. Provided clinical samples and data: P.D., R.G., G.P.K., R.G., S.S., K.N., R.K., S.N., V.K. Sample storage and record keeping: R.A., A.S. Clinical data management and analysis: R.A., A.B. Sample processing: R.A. Method standardization: R.A., A.S. Data acquisition: R.A., A.B. Data analysis: R.A. Investigation: R.A. Supervision: R.N., M.J.S Funding acquisition: R.N., R.G., R.K., S.N., G.P.K., R.G., S.S., K.N., M.J.S. Coordination and strategy: R.N., M.J.S. Writing-original draft: R.A. and R.N. Writing-review and editing: R.A., R.N., R.G., R.K., S.N., G.P.K., R.G., S.S., K.N., M.J.S. All authors read and approved the final manuscript and had full access to all the data in the study.

ETHICAL APPROVAL

The study was approved by the Institutional Ethics Committees of All India Institute of Medical Sciences (IEC-1074/06.11.20, RP-28/2020), Vardhman Mahavir Medical College (VMMC) and Safdarjung Hospital, New Delhi (IEC/VMMC/SJH/Project/2021-05/CC-157, dated 24.07.2021), Dr. Baba Saheb Ambedkar Medical College and Hospital (5(32)/2020/BSAH/DNB/PF, dated 31.08.2020), Lady Hardinge Medical College (LHMC/IEC/2022/03/30), University College of Medical Sciences & Guru Teg Bahadur Hospital (IECHC-2022-51-R1, dated 07.12.2022) and International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi (ICGEB/IEC/2021/28, version 2).

REFERENCES

1. K. E. Rudd, S. C. Johnson, K. M. Agesa, K. A. Shackelford, D. Tsoi, D. R. Kievlan, D. V. Colombara, K. S. Ikuta, N. Kissoon, S. Finfer, C. Fleischmann-Struzek, F. R. Machado, K. K. Reinhart, K. Rowan, C. W. Seymour, R. S. Watson, T. E. West, F. Marinho, S. I. Hay, R. Lozano, A. D. Lopez, D. C. Angus, C. J. L. Murray, M. Naghavi, Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *The Lancet* **395**, 200–211 (2020).
2. C. Fleischmann-Struzek, D. M. Goldfarb, P. Schlattmann, L. J. Schlapbach, K. Reinhart, N. Kissoon, The global burden of paediatric and neonatal sepsis: a systematic review. *Lancet Respir. Med.* **6**, 223–230 (2018).
3. I. Özmeral Odabaşı, Neonatal Sepsis. *SiSli Etfal Hastan. Tip Bul. Med. Bull. Sisli Hosp.*, doi: 10.14744/SEMB.2020.00236 (2020).
4. M. Singer, C. S. Deutschman, C. W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, R. Bellomo, G. R. Bernard, J.-D. Chiche, C. M. Coopersmith, R. S. Hotchkiss, M. M. Levy, J. C. Marshall, G. S. Martin, S. M. Opal, G. D. Rubenfeld, T. van der Poll, J.-L. Vincent, D. C. Angus, The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* **315**, 801–810 (2016).
5. Y. Li, J. Guo, H. Yang, H. Li, Y. Shen, D. Zhang, Comparison of culture-negative and culture-positive sepsis or septic shock: a systematic review and meta-analysis. *Crit. Care* **25**, 167 (2021).
6. D. Podlesny, W. F. Fricke, Strain inheritance and neonatal gut microbiota development: A meta-analysis. *Int. J. Med. Microbiol.* **311**, 151483 (2021).
7. B. M. Tang, G. D. Eslick, J. C. Craig, A. S. McLean, Accuracy of procalcitonin for sepsis diagnosis in critically ill patients: systematic review and meta-analysis. *Lancet Infect. Dis.* **7**, 210–217 (2007).
8. S. M. Yentis, N. Soni, J. Sheldon, C-reactive protein as an indicator of resolution of sepsis in the intensive care unit. *Intensive Care Med.* **21**, 602–605 (1995).
9. S. W. Wright, L. Lovelace-Macon, V. Hantrakun, K. E. Rudd, P. Teparrukkul, S. Kosamo, W. C. Liles, D. Limmathurotsakul, T. E. West, sTREM-1 predicts mortality in hospitalized patients with infection in a tropical, middle-income country. *BMC Med.* **18**, 159 (2020).
10. M. E. Mikkelsen, A. N. Miltiades, D. F. Gaieski, M. Goyal, B. D. Fuchs, C. V. Shah, S. L. Bellamy, J. D. Christie, Serum lactate is associated with mortality in severe sepsis independent of organ failure and shock*. *Crit. Care Med.* **37**, 1670 (2009).
11. S.-S. Huang, J.-Y. Lin, W.-S. Chen, M.-H. Liu, C.-W. Cheng, M.-L. Cheng, C.-H. Wang, Phenylalanine- and leucine-defined metabolic types identify high mortality risk in patients with severe infection. *Int. J. Infect. Dis.* **85**, 143–149 (2019).
12. T. C. Horan, M. Andrus, M. A. Dudeck, CDC/NHSN surveillance definition of health care–associated infection and criteria for specific types of infections in the acute care setting. *Am. J. Infect. Control* **36**, 309–332 (2008).
13. ICSH recommendations for measurement of erythrocyte sedimentation rate. International Council for Standardization in Haematology (Expert Panel on Blood Rheology). *J. Clin. Pathol.* **46**, 198–203 (1993).
14. J. Chong, D. S. Wishart, J. Xia, Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr. Protoc. Bioinforma.* **68**, e86 (2019).
15. B.-L. Grigorescu, Dubito Ergo Sum. Pathologies that can Mimic Sepsis. *J. Crit. Care Med.* **8**, 77–79 (2022).
16. W. Chen, W. Guo, Y. Li, M. Chen, Integrative analysis of metabolomics and transcriptomics to uncover biomarkers in sepsis. *Sci. Rep.* **14**, 9676 (2024).

17. R. Batra, R. Uni, O. M. Akchurin, S. Alvarez-Mulett, L. G. Gómez-Escobar, E. Patino, K. L. Hoffman, W. Simmons, W. Whalen, K. Chetnik, M. Buyukozkan, E. Benedetti, K. Suhre, E. Schenck, S. J. Cho, A. M. K. Choi, F. Schmidt, M. E. Choi, J. Krumsiek, Urine-based multi-omic comparative analysis of COVID-19 and bacterial sepsis-induced ARDS. *Mol. Med.* **29**, 13 (2023).
18. S. M. Gordon, L. Srinivasan, D. M. Taylor, S. R. Master, M. A. Tremoglie, A. Hankeova, D. D. Flannery, S. Abbasi, J. C. Fitzgerald, M. C. Harris, Derivation of a metabolic signature associated with bacterial meningitis in infants. *Pediatr. Res.* **88**, 184–191 (2020).
19. P. Liu, X. Zhang, X. Wang, Y. Liang, N. Wei, Z. Xiao, T. Li, R. Zhe, W. Zhao, S. Fan, Maternal sepsis in pregnancy and the puerperal periods: a cross-sectional study. *Front. Med.* **10** (2023).
20. S. M. Lee, M. Chang, K.-S. Kim, Blood Culture Proven Early Onset Sepsis and Late Onset Sepsis in Very-Low-Birth-Weight Infants in Korea. *J. Korean Med. Sci.* **30**, S67–S74 (2015).
21. F. R. Helmo, E. A. R. Alves, R. A. D. A. Moreira, V. O. Severino, L. P. Rocha, M. L. G. D. R. Monteiro, M. A. D. Reis, R. M. Etchebehere, J. R. Machado, R. R. M. Corrêa, Intrauterine infection, immune system and premature birth. *J. Matern. Fetal Neonatal Med.* **31**, 1227–1233 (2018).
22. Characterisation and antimicrobial resistance of sepsis pathogens in neonates born in tertiary care centres in Delhi, India: a cohort study. *Lancet Glob. Health* **4**, e752–e760 (2016).
23. K. M. Zangwill, A. Schuchat, J. D. Wenger, Group B Streptococcal Disease in the United States, 1990: Report from a Multistate Active Surveillance System. *Morb. Mortal. Wkly. Rep. Surveill. Summ.* **41**, 25–32 (1992).
24. P. T. Heath, G. Balfour, A. M. Weisner, A. Efstratiou, T. L. Lamagni, H. Tighe, L. A. O'Connell, M. Cafferkey, N. Q. Verlander, A. Nicoll, A. C. McCartney, Group B streptococcal disease in UK and Irish infants younger than 90 days. *The Lancet* **363**, 292–294 (2004).
25. M. T. Neto, Group B streptococcal disease in Portuguese infants younger than 90 days. *Arch. Dis. Child. - Fetal Neonatal Ed.* **93**, F90–F93 (2008).
26. Q. Chen, X. Liang, T. Wu, J. Jiang, Y. Jiang, S. Zhang, Y. Ruan, H. Zhang, C. Zhang, P. Chen, Y. Lv, J. Xin, D. Shi, X. Chen, J. Li, Y. Xu, Integrative analysis of metabolomics and proteomics reveals amino acid metabolism disorder in sepsis. *J. Transl. Med.* **20**, 123 (2022).
27. K. Feng, W. Dai, L. Liu, S. Li, Y. Gou, Z. Chen, G. Chen, X. Fu, Identification of biomarkers and the mechanisms of multiple trauma complicated with sepsis using metabolomics. *Front. Public Health* **10**, 923170 (2022).
28. M. Mierzchala-Pasierb, M. Lipinska-Gediga, M. G. Fleszar, P. Lesnik, S. Placzkowska, P. Serek, J. Wisniewski, A. Gamian, M. Krzystek-Korpacka, Altered profiles of serum amino acids in patients with sepsis and septic shock – Preliminary findings. *Arch. Biochem. Biophys.* **691**, 108508 (2020).
29. W. Druml, G. Heinzel, G. Kleinberger, Amino acid kinetics in patients with sepsis123. *Am. J. Clin. Nutr.* **73**, 908–913 (2001).
30. V. Fanos, P. Caboni, G. Corsello, M. Stronati, D. Gazzolo, A. Noto, M. Lussu, A. Dessì, M. Giuffrè, S. Lacerenza, F. Serraino, F. Garofoli, L. D. Serpero, B. Liori, R. Carboni, L. Atzori, Urinary 1H-NMR and GC-MS metabolomics predicts early and late onset neonatal sepsis. *Early Hum. Dev.* **90**, S78–S83 (2014).
31. B. P. Yoseph, N. J. Klingensmith, Z. Liang, E. R. Breed, E. M. Burd, R. Mittal, J. A. Dominguez, B. Petrie, M. L. Ford, C. M. Coopersmith, Mechanisms of intestinal barrier dysfunction in sepsis. *Shock Augusta Ga* **46**, 52–59 (2016).
32. G. Neurauter, A. V. Grahmann, M. Klieber, A. Zeimet, M. Ledochowski, B. Sperner-Unterweger, D. Fuchs, Serum phenylalanine concentrations in patients with ovarian

carcinoma correlate with concentrations of immune activation markers and of isoprostane-8. *Cancer Lett.* **272**, 141–147 (2008).
33. C. Jin, J. Ye, J. Yang, C. Gao, H. Yan, H. Li, X. Wang, mTORC1 Mediates Lysine-Induced Satellite Cell Activation to Promote Skeletal Muscle Growth. *Cells* **8**, 1549 (2019).
34. Y. Cao, Z. Wang, T. Yu, Y. Zhang, Z. Wang, Z. Lu, W. Lu, J. Yu, Sepsis induces muscle atrophy by inhibiting proliferation and promoting apoptosis via PLK1-AKT signalling. *J. Cell. Mol. Med.* **25**, 9724–9739 (2021).
35. K. Sarafidis, A. C. Chatziioannou, A. Thomaidou, H. Gika, E. Mikros, D. Benaki, E. Diamanti, C. Agakidis, N. Raikos, V. Drossou, G. Theodoridis, Urine metabolomics in neonates with late-onset sepsis in a case-control study. *Sci. Rep.* **7**, 45506 (2017).
36. K. A. Stringer, N. J. Serkova, A. Karnovsky, K. Guire, R. Paine, T. J. Standiford, Metabolic consequences of sepsis-induced acute lung injury revealed by plasma 1H-nuclear magnetic resonance quantitative metabolomics and computational analysis. *Am. J. Physiol.-Lung Cell. Mol. Physiol.* **300**, L4–L11 (2011).
37. X. Chen, B. Zhang, H. Li, X. Peng, Myo-inositol improves the host's ability to eliminate balofloxacin-resistant Escherichia coli. *Sci. Rep.* **5**, 10720 (2015).
38. R. An, L. Zhao, C. Xi, H. Li, G. Shen, H. Liu, S. Zhang, L. Sun, Melatonin attenuates sepsis-induced cardiac dysfunction via a PI3K/Akt-dependent mechanism. *Basic Res. Cardiol.* **111**, 8 (2015).

39. W. Zhong, K. Qian, J. Xiong, K. Ma, A. Wang, Y. Zou, Curcumin alleviates lipopolysaccharide induced sepsis and liver failure by suppression of oxidative stress-related inflammation via PI3K/AKT and NF-κB related signaling. *Biomed. Pharmacother.* **83**, 302–313 (2016).
40. J. B. Buse, J. L. R. Freeman, S. V. Edelman, L. Jovanovic, J. B. McGill, Serum 1,5-Anhydroglucitol (GlycoMark^{TM}): A Short-Term Glycemic Marker. *Diabetes Technol. Ther.* **5**, 355–363 (2003).
41. W. J. Kim, C.-Y. Park, 1,5-Anhydroglucitol in diabetes mellitus. *Endocrine* **43**, 33–40 (2013).
42. E. Yefet, S. Twafra, N. Shwartz, N. Hissin, J. Hasanein, R. Colodner, N. Mirsky, Z. Nachum, Inverse association between 1,5-anhydroglucitol and neonatal diabetic complications. *Endocrine* **66**, 210–219 (2019).
43. X. Meng, S. Tancharoen, K.-I. Kawahara, Y. Nawa, S. Taniguchi, T. Hashiguchi, I. Maruyama, 1,5-Anhydroglucitol Attenuates Cytokine Release and Protects Mice with Type 2 Diabetes from Inflammatory Reactions. *Int. J. Immunopathol. Pharmacol.* **23**, 105–119 (2010).
44. T. Gauthier, C. Yao, T. Dowdy, W. Jin, Y.-J. Lim, L. C. Patiño, N. Liu, S. I. Ohlemacher, A. Bynum, R. Kazmi, C. A. Bewley, M. Mitrovic, D. Martin, R. J. Morell, M. Eckhaus, M. Larion, R. Tussiwand, J. J. O'Shea, W. Chen, TGF-β uncouples glycolysis and inflammation in macrophages and controls survival during sepsis. *Sci. Signal.* **16**, eade0385 (2023).
45. M. Garcia-Alvarez, P. Marik, R. Bellomo, Sepsis-associated hyperlactatemia. *Crit. Care* **18**, 503 (2014).
46. E. M. Palsson-McDermott, L. A. J. O'Neill, The Warburg effect then and now: From cancer to inflammatory diseases. *BioEssays* **35**, 965–973 (2013).
47. J. Zhang, J. Muri, G. Fitzgerald, T. Gorski, R. Gianni-Barrera, E. Masschelein, G. D'Hulst, P. Gilardoni, G. Turiel, Z. Fan, T. Wang, M. Planque, P. Carmeliet, L. Pellerin, C. Wolfrum, S.-M. Fendt, A. Banfi, C. Stockmann, I. Soro-Arnáiz, M. Kopf, K. De Bock, Endothelial Lactate Controls Muscle Regeneration from Ischemia by Inducing M2-like Macrophage Polarization. *Cell Metab.* **31**, 1136-1153.e7 (2020).

48. J. Zhu, M. Zhang, T. Han, H. Wu, Z. Xiao, S. Lin, C. Wang, F. Xu, Exploring the Biomarkers of Sepsis-Associated Encephalopathy (SAE): Metabolomics Evidence from Gas Chromatography-Mass Spectrometry. *BioMed Res. Int.* **2019**, e2612849 (2019).
49. E. A. Hanse, C. Ruan, M. Kachman, D. Wang, X. H. Lowman, A. Kelekar, Cytosolic malate dehydrogenase activity helps support glycolysis in actively proliferating cells and cancer. *Oncogene* **36**, 3915–3924 (2017).
50. H. Yang, L. Du, Z. Zhang, Potential biomarkers in septic shock besides lactate. *Exp. Biol. Med.* **245**, 1066–1072 (2020).
51. G. M. Tannahill, A. M. Curtis, J. Adamik, E. M. Palsson-McDermott, A. F. McGettrick, G. Goel, C. Frezza, N. J. Bernard, B. Kelly, N. H. Foley, L. Zheng, A. Gardet, Z. Tong, S. S. Jany, S. C. Corr, M. Haneklaus, B. E. Caffrey, K. Pierce, S. Walmsley, F. C. Beasley, E. Cummins, V. Nizet, M. Whyte, C. T. Taylor, H. Lin, S. L. Masters, E. Gottlieb, V. P. Kelly, C. Clish, P. E. Auron, R. J. Xavier, L. a. J. O'Neill, Succinate is an inflammatory signal that induces IL-1β through HIF-1α. *Nature* **496**, 238–242 (2013).
52. T. van der Poll, M. Shankar-Hari, W. J. Wiersinga, The immunology of sepsis. *Immunity* **54**, 2450–2464 (2021).

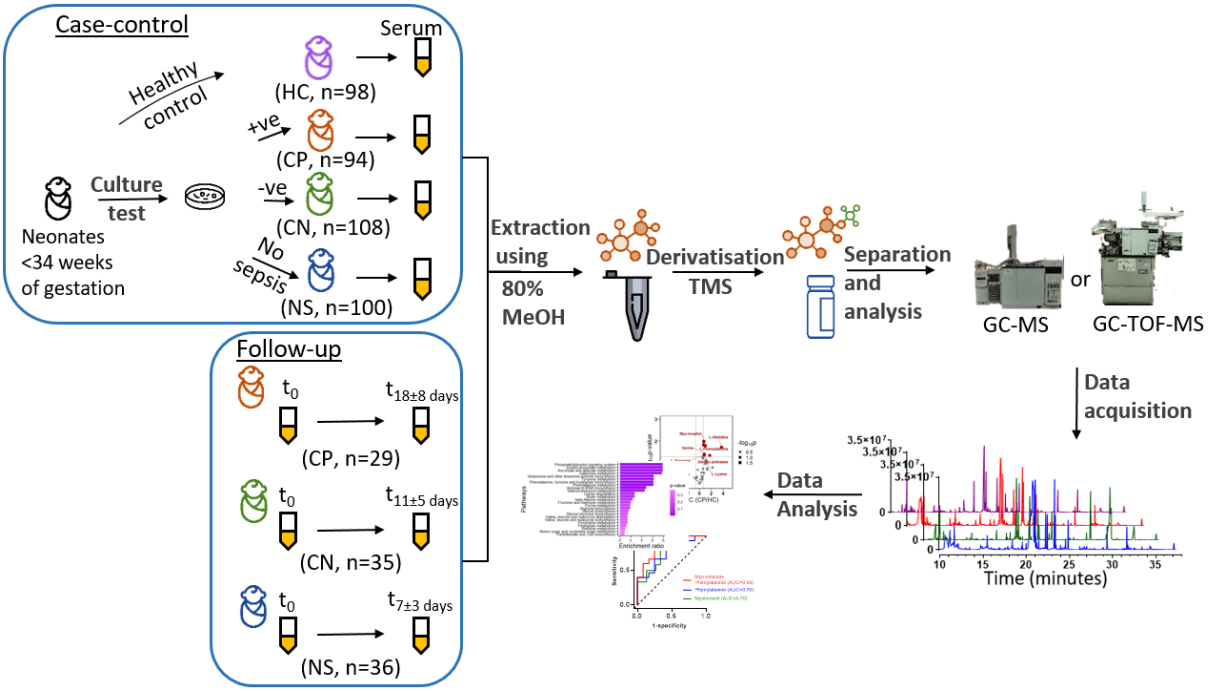

Figure 1: Schematic showing subject grouping, distribution, workflow used to identify the neonatal serum markers for sepsis. Blood culture was used to classify the subjects as sepsis (culture positive, CP; culture negative, CN) and no sepis (NS) group which was used as a control along with HC. Serum samples were used to extract metabolites for derivatization and mass spectrometry data acquisition and analysis using statistical tools. GC-MS: Gas Chromatography-Mass Spectrometry, GC-TOF-MS: Gas Chromatography-time-of-flight-Mass Spectrometry.

Table 1. Group-specific clinical epidemiological details of sepsis and control cases

Clinical details	Total	Discovery set-I					Discovery set-II					Validation set					Independent Follow-up set			
		Subtotal	Sepsis		Control		Subtotal	Sepsis		Control		Subtotal	Sepsis		Control		Subtotal	Sepsis		
			CP	CN	HC	NS		CP	CN	HC	NS		CP	CN	HC	NS		CP	CN	NS
Number of neonates	500	71	16	24	14	17	269	63	69	69	68	60	15	15	15	15	100	29	35	36
Gender (%, female)	41.2, 54	42.3, 30	50, 8	41.7, 10	42.9, 6	35.3, 6	43.4, 117	31.7, 20	49.3, 34	50.7, 35	41.1, 28	40, 24	60, 9	46.7, 7	20, 3	33.3, 5	46, 46	48.3, 14	40, 14	50, 18
Weight, mean±std. dev.* in kg	1.5 ±0.4	1.5 ±0.5	1.3 ±0.5	1.4 ±0.3	2 ±0.4	1.6 ±0.4	1.5 ±0.3	1.4 ±0.3	1.4 ±0.4	1.7 ±0.3	1.6 ±0.3	1.5 ±0.4	1.5 ±0.4	1.3 ±0.3	1.7 ±0.3	1.7 ±0.5	1.5 ±0.3	1.5 ±0.3	1.5 ±0.3	1.6 ±0.4
Mode of delivery (%, Caesarean)	30.5, 40	32.4, 23	31.3, 5	23.8, 5	43.8, 7	33.3, 6	73.9, 155	31.7, 20	71.8, 28	43.5, 30	51.5, 35	28.3, 17	40, 6	13.3, 2	26.7, 4	33.3, 5	46, 46	48.3, 14	51.4, 18	38.9, 14
Gestation, mean±std. dev.* in weeks	31.5 ±1.9	31.4 ±1.9	30.9 ±1.7	30.8 ±2.1	32.9 ±1.5	31.7 ±1.6	31.4 ±1.7	30.9 ±1.9	30.9 ±1.8	32.3 ± 1.3	31.6 ±1.5	31.6 ±1.9	30.9 ±2	30.5 ±1.9	32.6 ±1.2	32.3 ±1.5	31.5 ±1.8	31.2 ±2	31.2 ±1.7	32 ±1.6
Outcome (%, total)	23.7, 500	18.3, 71	43.8, 16	16.6, 24	-	11.8, 17	40.8, 250	77.3, 53	65, 63	1.5, 66	27.9, 68	30, 60	60, 15	60, 15	-	20, 15	-	-	-	-
Difficulty breathing (%, total)	65.2, 92	50, 52	43.8, 16	62.5, 22	-	33.3, 12	77.3, 181	81.1, 53	75.4, 61	-	76.1, 61	85, 40	100, 12	100, 13	-	33.3, 12	84, 99	86.2, 29	88.6, 35	80, 35
Apnea (%, total)	18.8, 92	23.1, 52	18.8, 16	27.3, 22	50, 2	8.3, 12	6.5, 182	13.2, 53	3.2, 63	-	4.5, 66	12.5, 40	25, 12	7.7, 13	-	6.7, 15	11, 100	20.7, 29	11.4, 35	2.8, 35
WBC count, mean±std. dev.* in $10^3/\mu L$	-	-	18.8 ±18	24.2 ±21.2	-	24.1 ±14.1	-	11.2 ±9.8	16.9 ±17.8	-	13.4 ±9.4	-	10.1 ±5	18.2 ±12.6	-	13.5 ±5.8	13.3 ±8.9	11.4 ±9	13.5 ±10.8	14.7 ±5.6
Absolute neutrophil count, mean±std. dev.* in 10^3	-	-	3.8 ±2.5	3.7 ±2.5	-	6.9 ±3.4	-	4.8 ±3.1	4.2 ±3.8	-	4.1 ±2.0	-	3.6 ±2.28	7.9 ±9.8	-	5.9 ±3.6	5.7 ±5.4	3.6 ±2.6	7.1 ±7.4	5.7 ±2.7

CP: culture-positive, CN: culture-negative, HC: healthy control, NS: no sepsis

*mean±standard deviation

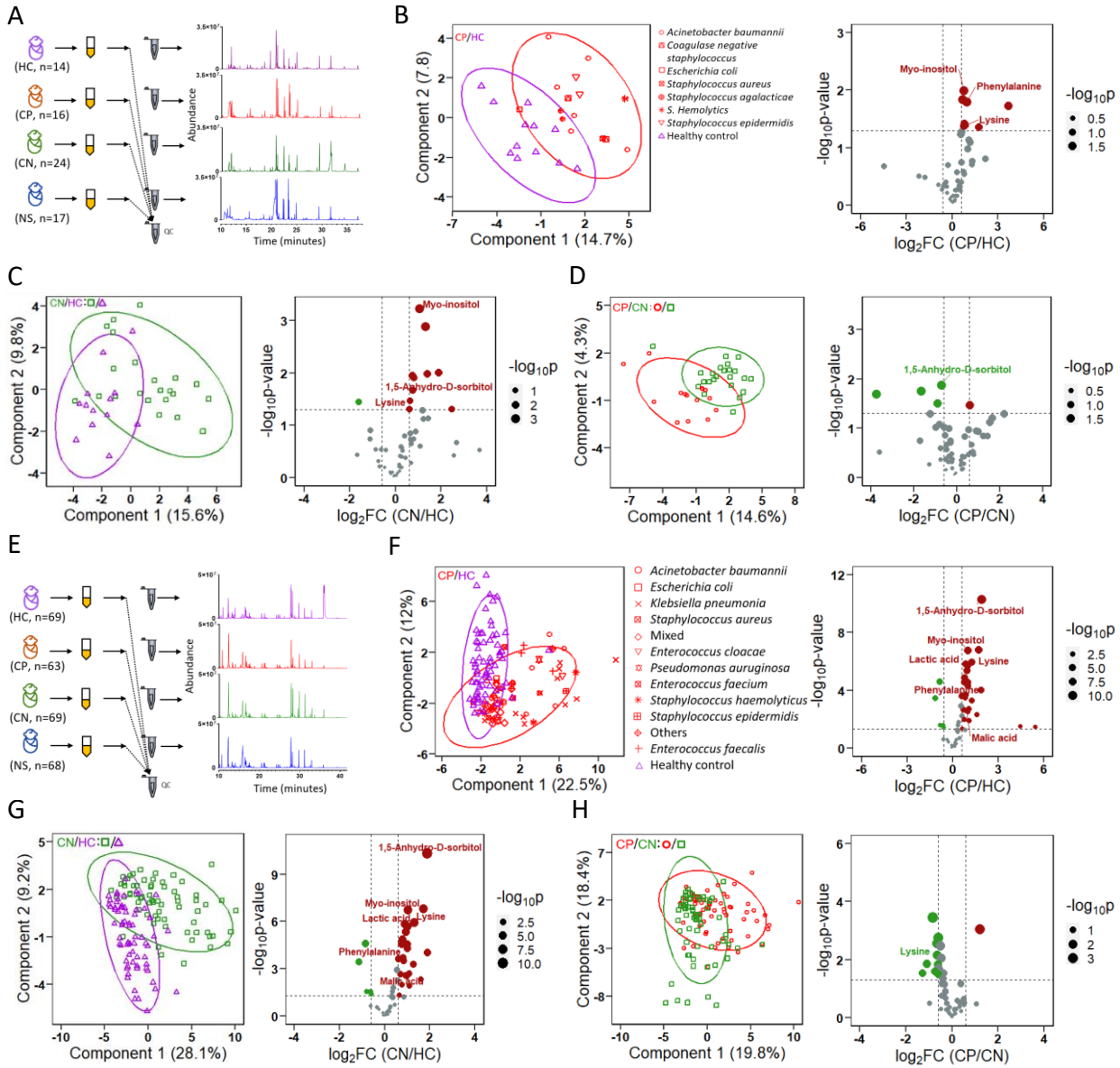

Figure 2: Serum metabolite profile of the discovery sets. (A) Representative Total ion chromatogram (TIC) of the serum metabolites harvested from the sepsis cases (culture positive/negative: CP/CN:16/24) and controls (no sepsis: NS:17 and healthy controls: HC:14) groups of discovery set-I (B-D) PLS-DA plots depicting the group-specific variations and volcano plots highlighting the significantly deregulated metabolic signatures ($p<0.05$, $FC>1.5$) between (B) CP and HC (C) CN and HC (D) CP and CN (A-D) Metabolite profile of discovery set-I acquired from quadrupole GC-MS (E) Representative Total ion chromatogram (TIC) of the serum metabolites harvested from the sepsis (CP/CN:63/69) cases and controls (NS/HC:68/69) groups of discovery set-II (F-H) PLS-DA plots depicting the group-specific variations and volcano plots highlighting the significantly deregulated metabolic signatures ($p<0.05$, $FC>1.5$) between (F) CP and HC (G) CN and HC (H) CP and CN (E-H) Metabolite profile of discovery set-II acquired from GC-TOF-MS (F) Volcano plots depicting the significantly deregulated metabolites between groups. PLS-DA: Partial Least Square-Discriminant Analysis, FC: Fold Change, GC-MS: Gas Chromatography-Mass Spectrometry, TOF: Time of Flight.

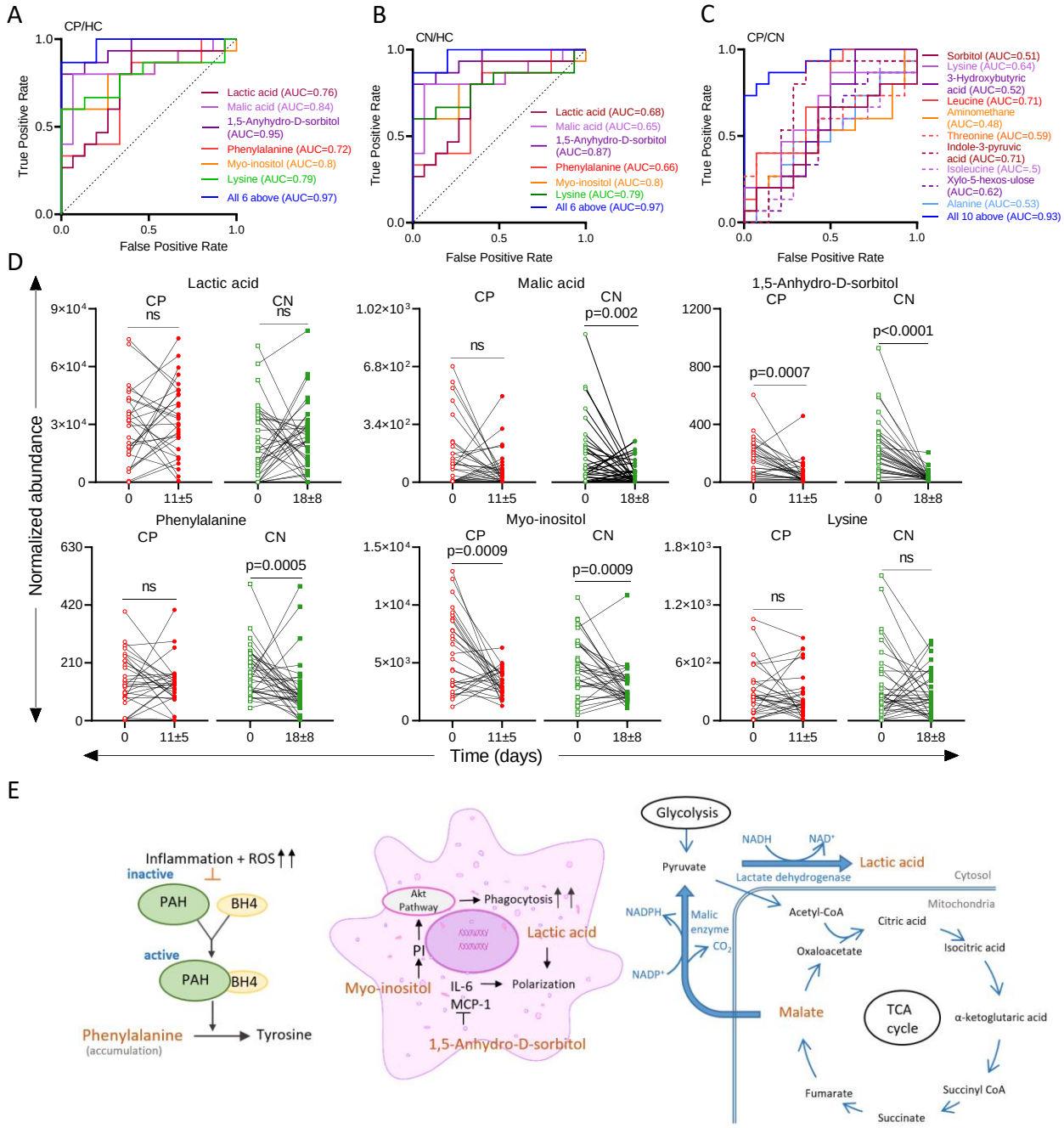

Figure 3: Validation of putative metabolic signature in independent data set and their trend in longitudinally followed up sepsis cases upon recovery. AUC of ROC of the deregulated set of metabolic features between (A) CP and HC group (B) CN and HC group (C) CP and CN group (CP/CN/HC:15 each) (D) Box plots depicting the change in abundance of individual deregulated metabolic features in longitudinally followed up neonatal sepsis cases upon recovery in an independent patient dataset (CP/CN:29/35) (E) Biological significance of the putative sepsis biomarkers. AUC of ROC: Area Under the Curve of Receiver Operating Characteristic, CP: Culture positive, CN: Culture negative, ns: not significant, PAH: Phenylalanine hydroxylase, BH4: Tetrahydrobiopterin, ROS: Reactive oxygen species, PI: Phosphatidyl-inositol, IL-6: Interleukin-6, MCP-1: Monocyte chemoattractant protein-1, NADH: Nicotinamide adenine dinucleotide, NADPH: Nicotinamide adenine dinucleotide phosphate. p-value <0.05 is considered significant.