DOI: 10.1002/jpen.2374

ORIGINAL RESEARCH

Fecal sphingolipids predict parenteral nutrition–associated cholestasis in the neonatal intensive care unit

Thomas J. Moutinho Jr PhD[1] | Deborah A. Powers MS[1] | Gabriel F. Hanson BA[1] | Shira Levy MA[2] | Rajiv Baveja MD[3] | Isabel Hefner MD[2] | Masouma Mohamed MD[2] | Alaa Abdelghani MD[2] | Robin L. Baker MD[3] | Jason A. Papin PhD[1] | Sean R. Moore MD, MS[4] | Suchitra K. Hourigan MD[2,4,5,6]

[1]Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, USA

[2]Inova Children's Hospital, Falls Church, Virginia, USA

[3]Fairfax Neonatal Associates, Falls Church, Virginia, USA

[4]Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of Virginia, Charlottesville, Virginia, USA

[5]Division of Pediatric Gastroenterology, Pediatric Specialists of Virginia, Fairfax, Virginia, USA

[6]Laboratory of Host Immunity and Microbiome, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA

Correspondence
Jason A. Papin, PhD, Department of Biomedical Engineering, University of Virginia, MR-5 Building, 415 Lane Rd, Charlottesville, VA 22908, USA.
Email: papin@virginia.edu

Sean R. Moore, MD, MS, Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Department of Pediatrics, University of Virginia, MR-4 Building, 409 Lane Rd, Charlottesville, VA 22908, USA.
Email: sean.moore@virginia.edu

Suchitra K. Hourigan, MD, Laboratory of Host Immunity and Microbiome, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 50, Room 5511, 50 South Dr, Bethesda, MD 20892, USA.
Email: suchitra.hourigan@nih.gov

Funding information
Division of Intramural Research, National Institute of Allergy and Infectious Diseases; National Center for Advancing Translational Sciences, Grant/Award Number: UL1TR003015; National Institute of Child Health and Human Development, Grant/Award Number: K23HD099240

Abstract

Background: Parenteral nutrition–associated cholestasis (PNAC) in the neonatal intensive care unit (NICU) causes significant morbidity and associated healthcare costs. Laboratory detection of PNAC currently relies on elevated serum conjugated bilirubin levels in the aftermath of impaired bile flow. Here, we sought to identify fecal biomarkers, which when integrated with clinical data, would better predict risk for developing PNAC.

Methods: Using untargeted metabolomics in 200 serial stool samples from 60 infants, we applied statistical and machine learning approaches to identify clinical features and metabolic biomarkers with the greatest associative potential for risk of developing PNAC. Stools were collected prospectively from infants receiving PN with soybean oil–based lipid emulsion at a level IV NICU.

Results: Low birth weight, extreme prematurity, longer duration of PN, and greater number of antibiotic courses were all risk factors for PNAC ($P < 0.05$). We identified 78 stool biomarkers with early predictive potential ($P < 0.05$). From these 78 biomarkers, we further identified 12 sphingomyelin lipids with high association for the development of PNAC in precholestasis stool samples when combined with birth anthropometry.

Conclusion: We demonstrate the potential for stool metabolomics to enhance early identification of PNAC risk. Earlier detection of high-risk infants would empower

proactive mitigation with alterations to PN for at-risk infants and optimization of energy nutrition with PN for infants at lower risk.

KEYWORDS
early detection, infant, metabolomics, neonatal intensive care unit, parenteral nutrition–associated cholestasis, sphingomyelin, stool

CLINICAL RELEVANCY STATEMENT

Parenteral nutrition–associated cholestasis (PNAC) in premature neonates is associated with significant morbidity and healthcare costs. Current detection of PNAC relies on elevated serum conjugated bilirubin levels after liver damage has already occurred. Here, we demonstrate fecal metabolomics enhances early identification of infants at risk of developing PNAC prior to elevations in serum conjugated bilirubin. Earlier detection would empower proactive mitigation of PNAC, perhaps by guiding selection of standard vs next-generation lipid formulations or optimization of PN for infants at lower risk of PNAC.

INTRODUCTION

Parenteral nutrition (PN) use is essential in the care of premature or ill neonates in the neonatal intensive care unit (NICU). A subset of infants who receive PN develop liver damage and PN-associated cholestasis (PNAC).[1] PNAC is diagnosed by an elevated level of serum conjugated bilirubin that can only be detected once there is decreased bile flow through the bile ducts and accumulation of bile products in the blood due to impaired excretion.[2] There are several recognized clinical risk factors for the development of PNAC, including longer duration of PN, low birth weight, and lower gestational age.[3,4] Recent improvements in feeding protocols have helped to lower this risk; however, the incidence of PNAC exceeds 85% in infants requiring PN for longer than 14 weeks. Additionally, a birth weight of <750 g, which is also associated with prolonged PN use, is predictive of PNAC.[5] Liver injury can persist even after cessation of PN, representing a significant health and economic burden.[6] Recently, alternative PN lipid emulsions sourced from fish (Omegaven) have been shown to limit the progression and injury from PNAC in a meta-analysis. There is also evidence of benefit for lipid emulsions sourced from a mixture of plant and fish lipids (Smoflipid); however, large randomized control trials have not yet been reported.[7,8] Becaues of multiple factors including cost, insurance limitations, and the requirement for a second intravenous catheter to administer incompatible drugs, these potentially protective lipid formulations are not used for all infants in the NICU.[9] Therefore, there is a great need for identifying effective biomarkers (early in life) to predict which infants will develop PNAC in the NICU, prior to a clinical diagnosis of PNAC, to be able to target hepatoprotective strategies in those most at risk.[10]

Stool and urine provide valuable information about physiological processes and their collection is noninvasive with minimal risk.[10,11] However, stool and urine remain an untapped resource for diagnostic applications in the NICU. We and others have demonstrated an association between the development of PNAC and a differential composition in the gut microbiome.[12–15] These findings suggest detectable changes in the stool might predict PNAC.

In this study, our aim was to identify the stool metabolomic and microbiome signatures predictive of PNAC and integrate this with clinical information to better predict which infants are at higher risk for developing PNAC. We hypothesized that metabolites present in the stool would present earlier than the current diagnostic metric, elevated serum conjugated bilirubin levels, thus providing early detection of infants at risk of PNAC.

METHODS

Sample collection and processing

Participants were chosen from a larger cohort of ongoing neonatal microbiome studies from the level IV NICU at Inova Fairfax Hospital in Virginia (Western Institutional Review Board approval number 20210065). Between 2016 and 2019, after informed consent, preterm and term neonates aged ≤5 days with an anticipated medical intermediate care unit length of stay of >5 days were recruited. We collected detailed maternal, pregnancy, and delivery data. While in the NICU, infants had stool collected twice weekly when possible. Stool was frozen at −80°C within 12 h. To address the aim of the current study, 60 infants were selected from the larger cohort who all had received PN for >5 days. Of these, all were receiving partial enteral nutrition at the time of the first stool sample collection. For stool metabolomics analyses, we enriched sample selection for PNAC cases and excluded samples in which PNAC was presumed to be secondary to infection. Infants were monitored for the development of cholestasis (defined as a conjugated bilirubin ≥ 1 mg/dl[16]), and the cause of cholestasis was noted as assessed by the treating physicians. Clinical data were collected at baseline and with each stool collection; this included gestational age, birth weight, antibiotic use, infections, and length of time receiving PN.

16S ribosomal RNA (rRNA) gene sequencing

Three hundred and twenty-seven samples underwent 16S rRNA gene sequencing of the hypervariable region V4, at either the Inova Core research laboratory or Ubiome. At the Inova Core research laboratory, DNA was extracted using the DNeasy PowerSoil Pro Kit (Qiagen)

following the manufacturer's protocol and sequenced on the Miseq platform (Illumina). Samples sent to Ubiome were sequenced following previously reported methods with 150 base pair paired-end reads.[17]

Metabolomics

Two hundred stool samples were prepared for metabolomic analysis performed, as previously described.[18] Briefly, frozen samples were lyophilized and then resuspended at a 50:1 (50 µl deionized water for every 1 mg of feces weight) ratio for homogenization.[19] The homogenates were subjected to automated biochemical extraction and analysis by liquid chromatography and high-resolution tandem mass spectrometry (MS/MS) on Metabolon's Global Platform.[20–22] Raw data were extracted, peak identified, and processed by Metabolon using proprietary software.[19,23,24] Briefly, metabolites were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a dynamic and proprietary biochemical reference library of >4500 known metabolites (based on authenticated standards) and >2000 novel metabolites (without an identified chemical structure); each library entry contains the retention time/index (RI), mass-to-charge ratio, and spectral data (including MS/MS fragmentation). Biochemical identifications are based on three criteria: RI within a narrow RI window of the proposed identification, accurate mass match to the library ± 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. Three types of controls were included: a pool of small portions of each experimental sample, serving as a technical replicate throughout the platform run; extracted water samples (process blanks); and a mixture of standards spiked into every analyzed sample, allowing instrument performance monitoring.

Statistical analysis

We analyzed the clinical metadata for each infant enrolled in the study as well as the metabolomics data for each of the fecal samples. For clinical data, Mann-Whitney *U* tests were performed to correlate clinical metrics with conjugated bilirubin levels. For biomarker identification, if metabolites were correlated with conjugated bilirubin levels, Mann-Whitney *U* tests were performed followed by a multiple testing correction using the Bonferroni method to identify metabolites differentially abundant in healthy vs control groups. Distributions were confirmed to be nonnormal using the Shapiro-Wilk test.

Data processing and machine learning

We used both QIIME and DADA2 software to process the 16S data.[25–27] The forward reads from the 16S data were used to align with the Greengenes 16S rRNA gene database to determine taxonomic calls. We processed the metabolomics data using Python, Jupyter notebooks, pandas, scipy, numpy, and seaborn. We generated the machine learning models using Scikit Learn in Python. All Python scripts and data are available on github (https://github.com/Tjmoutinho/TPNAC_data_processing).

RESULTS

This analysis included 60 infants, all of whom received PN with Intralipid 20% (Baxter) as the lipid emulsion. All infants were also receiving some enteral nutrition at the time of their first stool sample was analyzed. During the study, 19 of 60 (32%) infants developed PNAC (see Table S1 for clinical and demographic characteristics of the cohort stratified by whether the infant developed PNAC). The cholestasis of these infants was not attributed to any other cause. In total, 391 stool samples were collected over the course of the study. Using these samples, we performed 16S rRNA gene sequencing ($n = 327$) to characterize the intestinal microbiome and conducted liquid chromatography–mass spectrometry ($n = 200$) to measure the stool metabolome.

Clinical metrics can identify an at-risk population of infants receiving PN

We collected relevant clinical data for all infants enrolled in this study, including but not limited to their gestational age at birth, birth weight, number of days of PN before developing PNAC, antibiotic use, and calculated metrics such as birth weight percentile adjusted for gestational age (Table S1). Among the clinical metrics measured, only birth weight and the number of days of PN before occurrence of PNAC were statistically different between the disease and control groups ($P < 0.05$) (Figure 1A). We found that infants with a birth weight percentile of >40% and a birth weight of >1.1 kg were less likely to develop PNAC. These clinical criteria are consistent with known clinical risk factors (Figure 1B). Since length of PN administration is associated with the development of PNAC, we investigated the relationship further in our cohort to better understand how PNAC is impacted by duration of PN. We observed a positive correlation between PNAC diagnosis and the amount of time an infant receives PN before diagnosis ($P < 0.05$). Two-thirds of the infants diagnosed with PNAC received PN for longer than 20 days. However, the length of time receiving PN is limited as an early clinical predictor of the risk of development of PNAC, as it is not concretely known until later in the clinical course, although certain clinical scenarios predict a longer length of PN.

Bacterial taxa correlate with PNAC but lack predictive potential in this cohort

Within this cohort of NICU infants, we identified several microbial taxa statistically different between the disease and control groups ($P < 0.05$). Among these taxa, the Enterobacteriaceae were present at greater abundance in stool samples collected from infants with serum

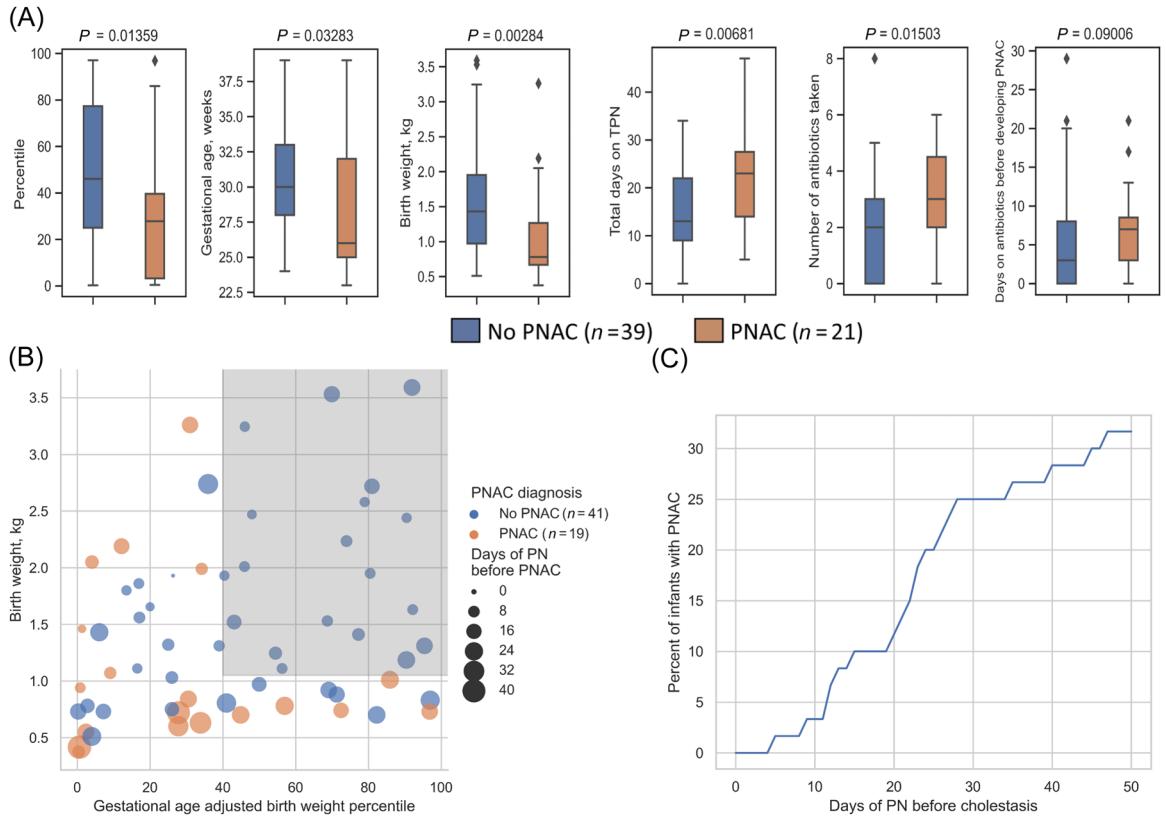

FIGURE 1 Clinical characteristics of infants with and without parenteral nutrition–associated cholestasis (PNAC). (A) Continuous clinical variables were tested for a statistically significant difference between the control and PNAC groups using a Wilcoxon rank sum test. We determined that five of the six variables are statistically different ($P < 0.05$). (B) The comparison of two clinical metrics, birth weight and birth weight percentile, reveals that there are simple thresholds that classify infants in our cohort as high or low risk. Infants born above the 40th weight percentile (adjusted for gestational age) and also weigh >1.1 kg at birth are at a low risk of developing PNAC compared with the rest of the infants. (C) Although PNAC diagnosis is correlated with the amount of time an infant receives PN (panel A), two-thirds of infants diagnosed with PNAC received PN for >20 days.

conjugated bilirubin levels ≥ 1 mg/dl (Figure S1A). However, the three known genera within the Enterobacteriaceae family all demonstrate an opposite trend, with high abundance in the control group (Figure S1B). The unknown genera within the Enterobacteriaceae family showed a higher abundance in samples from infants with PNAC (Figure S1C). Although this result is consistent with our previous work,[14] more research is needed to identify the primary Enterobacteriaceae genera present at elevated abundance in infants with PNAC. Finally, at the species level, *Veillonella dispar* was at an elevated relative abundance in infants with PNAC (Figure S1D). It is important to note that none of these microbes showed strong associations for determining infants at risk of developing PNAC; all are simply correlated with elevated levels of conjugated bilirubin.

The stool metabolome contains valuable biomarkers for infants at risk of PNAC

Among the 19 of 60 infants in our cohort who developed PNAC (Figure 2A), there were nine for whom we were able to collect stool samples from before a diagnosis of PNAC. The interval between fecal metabolomic measurements and the initial detection of PNAC varied based on timing of sample collection. The timing of stool sample measurements relative to the detection of PNAC in 19 patients is shown in Figure 2A. In the nine children (indicated by the dotted blue circles in Figure 2A) from whom we collected ≥1 fecal samples before PNAC was diagnosed, the interval between measured elevations in fecal sphingomyelins and the detection of PNAC ranged from 5 to 46 days, with a mean and median of 20.6 and 15 days, respectively. These early samples of stools from children who went on to develop PNAC were used as a filter to provide a glimpse into metabolites with the greatest association for the risk of developing PNAC prior to disease progression. We first identified >100 metabolites differentially abundant between the disease group and the control group. For each candidate metabolite identified, we then compared the scaled quantity of that metabolite in each of the nine case study samples to the median value in all of the control samples. For metabolites found to be more abundant in PNAC samples, we selected only those with scaled intensities in >90% of the nine case study samples above the median value of the PNAC distribution (Figure 2B). For metabolites negatively correlated with disease, we selected PNAC biomarkers by identifying those with scaled intensities in >90% of the nine case study samples falling below the median value of the disease

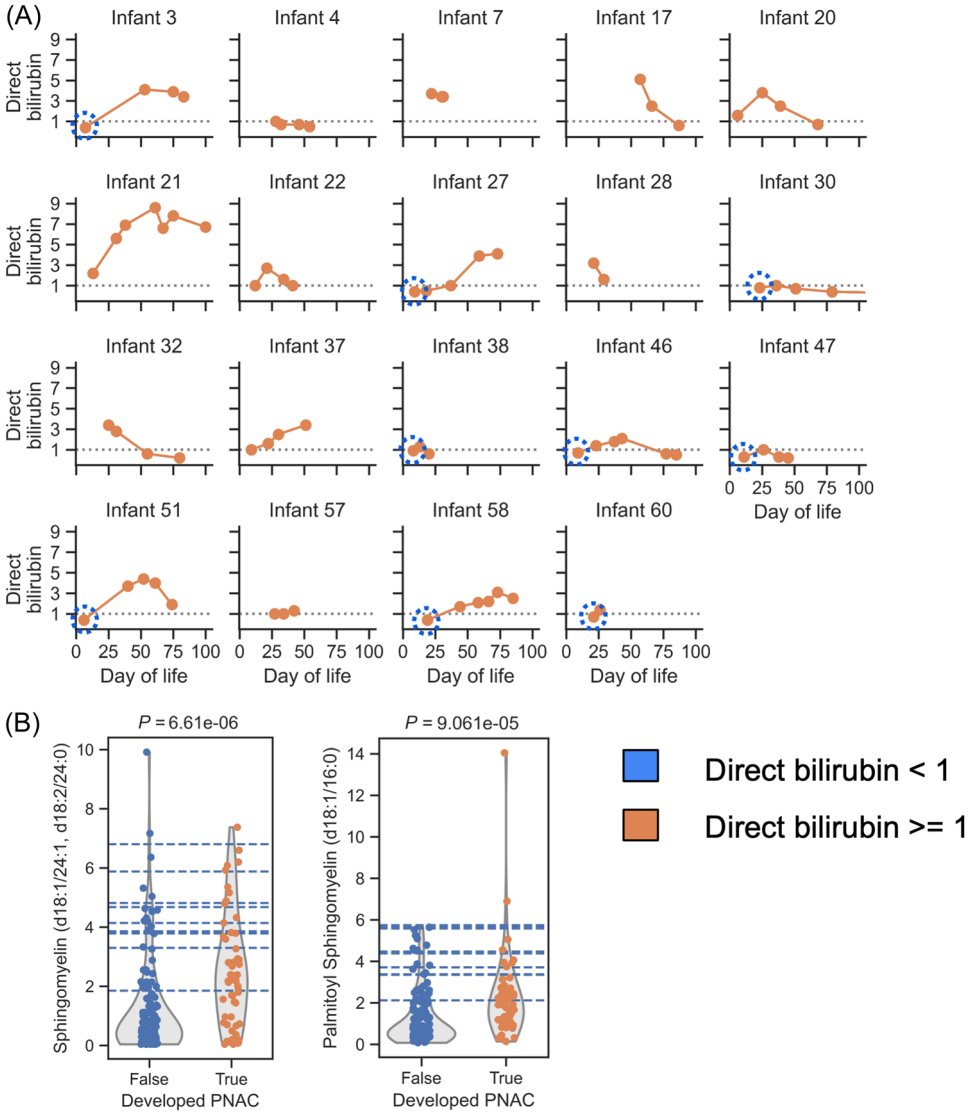

FIGURE 2 Metabolomics data and predictive biomarker selection. (A) There are 60 infants in this study and 19 were diagnosed with parenteral nutrition–associated cholestasis (PNAC). We collected 200 fecal samples with accompanied metabolomics data. The samples from each infant are plotted on an individual plot. The clinical conjugated bilirubin threshold, used to diagnose PNAC, is displayed with a dashed gray line on each panel. There are nine cholestatic infants for which we were able to collect fecal samples before conjugated bilirubin levels were above the diagnostic threshold of 1 mg/dl, as indicated by the dashed blue circle. (B) The y-axis displays the scaled intensity for each metabolite. The case study samples were plotted as dashed blue lines across the boxplots. We selected biomarkers based on statistical significance ($P < 0.05$) and consistency among the case study samples.

distribution. We performed Mann-Whitney *U* tests to assess the statistical significance of each metabolite between the samples with conjugated bilirubin below the clinical threshold and above. We used a Bonferroni multiple comparison correction to control for the number of metabolites tested. This analysis resulted in the identification of a subset of metabolites with the greatest potential to be early predictors for the risk of developing PNAC (Table 1).

After multiple testing corrections, we identified a total of 45 biomarkers identified in early stool samples associated with the risk of developing PNAC. Specifically, we found 25 biomarkers positively correlated with a subsequent rise in conjugated bilirubin levels (Table 1) and 20 metabolites negatively correlated with a subsequent increase in conjugated bilirubin levels (Figure S2). The biomarkers elevated in PNAC samples provide a glimpse into the pathophysiology that may be governing the disease. Although these biomarkers have early predictive value based on the case study infants, they require additional validation in future studies via targeted metabolomics.

Individual biomarkers are quite useful; however, a comparison of metabolites based on general classifications provides a measure of uniqueness for each biomarker. For example, we identified 1000 known metabolites across all samples and a total of 19 sphingomyelin metabolites. We determined 18 of 19 known sphingomyelin metabolites were biomarkers for elevated serum conjugated bilirubin levels. Separately, we classified PNAC biomarkers into seven

TABLE 1 Biomarkers positively correlated with PNAC.

Metabolite	FDR-corrected *P* value
1-Linoleoyl-GPC (18:2)	1.86E−05
1-Oleoyl-2-linoleoyl-GPC (18:1/18:2)	8.63E−03
1-Oleoyl-GPC (18:1)	3.51E−05
1-Palmitoyl-GPC (16:0)	1.88E−02
1-Stearoyl-GPC (18:0)	6.01E−04
1-Stearoyl-GPE (18:0)	8.56E−07
2-Stearoyl-GPE (18:0)	4.11E−04
Behenoyl sphingomyelin (d18:1/22:0)	1.18E−05
Lignoceroyl sphingomyelin (d18:1/24:0)	6.22E−05
Palmitoyl dihydrosphingomyelin (d18:0/16:0)	1.87E−04
Palmitoyl sphingomyelin (d18:1/16:0)	9.06E−05
Sphingomyelin (d17:1/16:0, d18:1/15:0, d16:1/17:0)	1.24E−05
Sphingomyelin (d18:0/18:0, d19:0/17:0)	6.09E−04
Sphingomyelin (d18:1/14:0, d16:1/16:0)	5.02E−07
Sphingomyelin (d18:1/17:0, d17:1/18:0, d19:1/16:0)	1.29E−05
Sphingomyelin (d18:1/20:0, d16:1/22:0)	4.93E−06
Sphingomyelin (d18:1/20:1, d18:2/20:0)	5.71E−07
Sphingomyelin (d18:1/21:0, d17:1/22:0, d16:1/23:0)	7.77E−07
Sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1)	8.12E−07
Sphingomyelin (d18:1/24:1, d18:2/24:0)	6.61E−06
Sphingomyelin (d18:2/16:0, d18:1/16:1)	1.14E−03
Sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)	5.74E−05
Sphingomyelin (d18:2/24:1, d18:1/24:2)	1.77E−06
Stearoyl sphingomyelin (d18:1/18:0)	3.43E−07
Tricosanoyl sphingomyelin (d18:1/23:0)	1.73E−05

Note: We identified 32 biomarkers that are statistically significantly elevated in samples with an associated conjugated bilirubin level of ≥1 mg/dl, whereas all nine case study samples have values that are above the median value of the elevated group.

Abbreviations: FDR, false discovery rate; PNAC, parenteral nutrition–associated cholestasis.

additional groups (Table 2). Several other classifications of membrane lipids present at greater abundance in PNAC samples indicate a general dysregulation of lipid metabolism in the liver or gastrointestinal (GI) tract. Several primary and secondary bile acids were reduced in PNAC samples, validating well-known pathophysiology. Long-chain carnitines were also reduced in PNAC samples.

The sphingomyelin biomarkers presented in Tables 1 and 2 and Figure S2 are directly applicable to the clinical setting. They have the potential to predict the development of PNAC before clinical markers of cholestasis rise, that is, diagnosing PNAC after detecting elevations in serum conjugated bilirubin levels.

Birth weight and birth weight percentile (adjusted for gestational age) were the most important predictors of infants at risk of developing PNAC (Figure 3A). When including metabolic biomarkers, we were able to achieve >70% overall predictive accuracy, with five fold cross-validation when classifying samples as control or PNAC (Figure 3B). Five of the top six biomarkers determined to be the most predictive were sphingomyelins, as were 16 of the top 20 predictive biomarkers. Birth weight percentile and birth weight, again, demonstrate their predictive utility while accounting for the metabolic biomarkers. Other important biomarkers in our random forest models include several long-chain carnitines and bile acids.

A small set of biomarkers and clinical variables provide early indication of infants at risk of PNAC

With predictive potential demonstrated using fivefold cross-validation with a set of random forest machine learning models, we next calculated the classification accuracy within our cohort using simple criteria that would be possible to implement in the clinic. The ideal implementation of these biomarkers in the clinic would be a simple point-of-care diagnostic that provides neonatologists with additional information regarding an infant's gut health. This ideal scenario requires basic thresholds to be applied to a small set of metabolites present in the stool in the first weeks of life recorded at birth, which provide valuable PNAC risk stratification. In our cohort, infants with birth weights >40th percentile for gestational age and >1.1 kg were at a significantly lower risk for developing PNAC. We assigned the other 41 infants who did not meet thesec cutoffs in a high-risk group to determine which fecal biomarkers provided the most robust discriminatory accuracy for the development of PNAC (Figure 4). We found the 12 best biomarkers, selected from a complete set of 78, were all sphingomyelin metabolites. Each of these metabolites provided a classification accuracy of the infants in our study between 82% and 88%, superior to the 66% classification accuracy provided by anthropometrics alone.

Although the 12 metabolites individually discriminated between the disease and control groups in our cohort, it was important to not overfit our model and apply a more robust strategy focused on future implementation in the NICU. To improve the robustness of our calculations, we assessed the accuracy of classification when using all 12 metabolites simultaneously by using an ensemble approach. For this, we analyzed the first available stool samples from infants meeting the following criteria: (1) high risk by clinical metrics and (2) stool samples collected within the first 3 weeks of life. For the 34 samples meeting these criteria, a majority vote ensemble classifier for our cohort provides an overall accuracy of 85%, which demonstrates consistency in the alignment across the 12 metabolites identified as optimal candidates for a rigorous follow-up validation study (Figure 5).

TABLE 2 Biomarker classifications.

	Biomarker classification	Number of biomarkers	Total in database	Biomarker group statistically significant, %	Positively or negatively correlated with conjugated bilirubin levels
1	Sphingomyelins	18	19	95	positive
2	GPC phosphatidylcholines	1	13	8	positive
3	GPC lysophospholipids	4	6	66	positive
4	GPE lyso-lipids	2	8	25	positive
5	Primary cholic bile acids	4	17	24	negative
6	Secondary cholic bile acids	4	32	13	negative
7	Vitamin A and carotene diols	2	4	50	negative
8	Long-chain carnitines (≥10 carbons)	10	28	36	negative
	Total statistically significant biomarkers	45	—	—	—
	Total identified metabolites in data set	—	1000	—	—

Notes: We classified each biomarker into 10 groups. Among all of the sphingomyelin molecules in the known set of metabolites detected, we identified all of them to have biomarker potential for early prediction of PNAC. Em dashes (—) denote information that was not applicable.

Abbreviation: PNAC, parenteral nutrition–associated cholestasis.

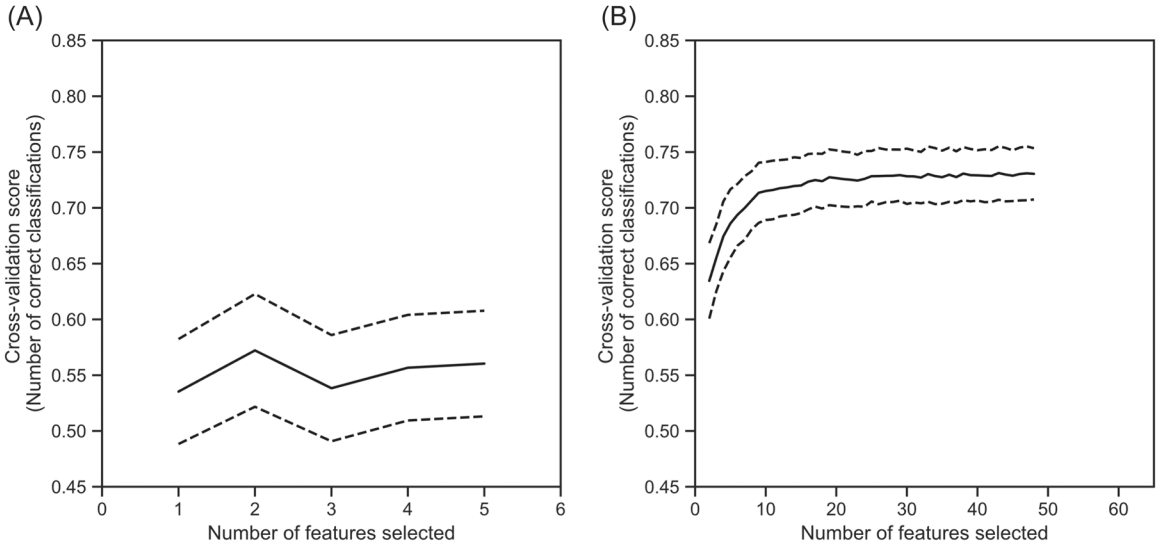

FIGURE 3 Random forest machine learning with fivefold cross-validation. (A) We performed feature reduction random forest machine learning to determine the minimal set of clinical metrics that provide the greatest predictive potential in our cohort. The optimal random forest consists of two clinical metrics and has an average fivefold cross-validation overall accuracy of 57%. Birth weight percentile adjusted by gestational age and birth weight each contribute significantly to this model. (B) When we include the 78 biomarkers that we identified to have predictive potential, we are able to generate a set of random forest models with >70% cross-validation accuracy on average. The second set of models also demonstrates that the two previously identified clinical variables maintain predictive potential when in the context of the stool biomarkers. These models use all of the metabolomics samples in this study and classify samples as high or low bilirubin levels; therefore, they do not predict if an infant will develop parenteral nutrition–associated cholestasis from the early stool samples collected.

DISCUSSION

In this NICU study of fecal biomarker predictors of PNAC in 200 fecal samples for 60 infants, our key novel finding is identification of 78 metabolites present in stool samples associated with an increased risk of an infant developing PNAC. Importantly, fecal metabolites, in conjunction with early anthropometry, provided greater predictive value than clinical factors alone. Notably, 12 sphingomyelin lipids demonstrated significant predictive potential in our cohort. Although the number of patients with stool samples before and after a diagnosis of PNAC is small ($n = 9$) and limits how far these results may be extrapolated, this analysis does identify promising metabolites to validate moving forward, with an ultimate goal of a diagnostic test for early prediction of risk of

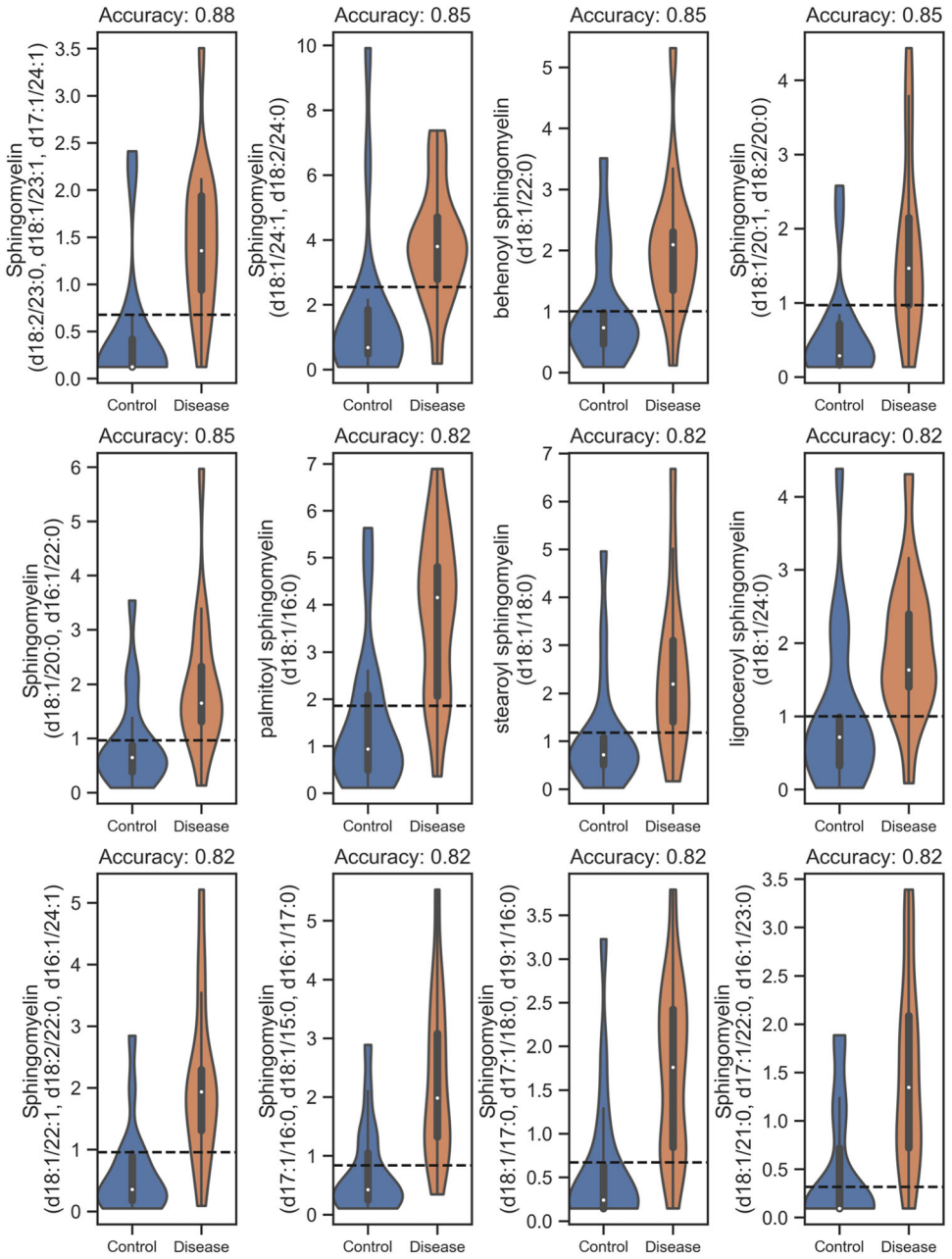

FIGURE 4 Biomarkers with the strongest discriminatory accuracy. This analysis includes only the infants who fall outside of the low-risk group identified in Figure 1, with a birth weight percentile of >40% and a birth weight of >1.1 kg. Additionally, we reduced the number of stool samples to only include the first sample for each infant. There are 12 metabolites from our complete set of 78 that demonstrate particularly accurate discriminatory potential within our cohort. These metabolites range from being 88% to 82% accurate at classifying the infants in our cohort based on only the first fecal sample that was collected for each infant. Although these accuracies are not properly validated with independent data, they demonstrate that there are several metabolites present in neonatal intensive care unit stool samples that have predictive capabilities. All 12 of these metabolites are various types of sphingomyelin.

developing PNAC. In addition, we again identified measurable differences in gut microbiota between infants with and without PNAC. Whether fecal sphingomyelins are mechanistically connected to these differences in gut microbiota remains to be seen. Further work is needed to establish the potential value of fecal sphingomyelins in informing precision clinical decision-making involving PN formulations and PNAC prevention.

PNAC is a common adverse outcome of lifesaving PN. Advancements in precision nutrition in the NICU may ultimately allow medical teams to improve nutrition plans and health outcomes. A key challenge in the NICU is the lack of access to frequent blood samples for longitudinal, predictive diagnostic testing. Infant stools, although not without logistical challenges, provide noninvasive, nonexhaustible samples with significant predictive potential in the NICU.

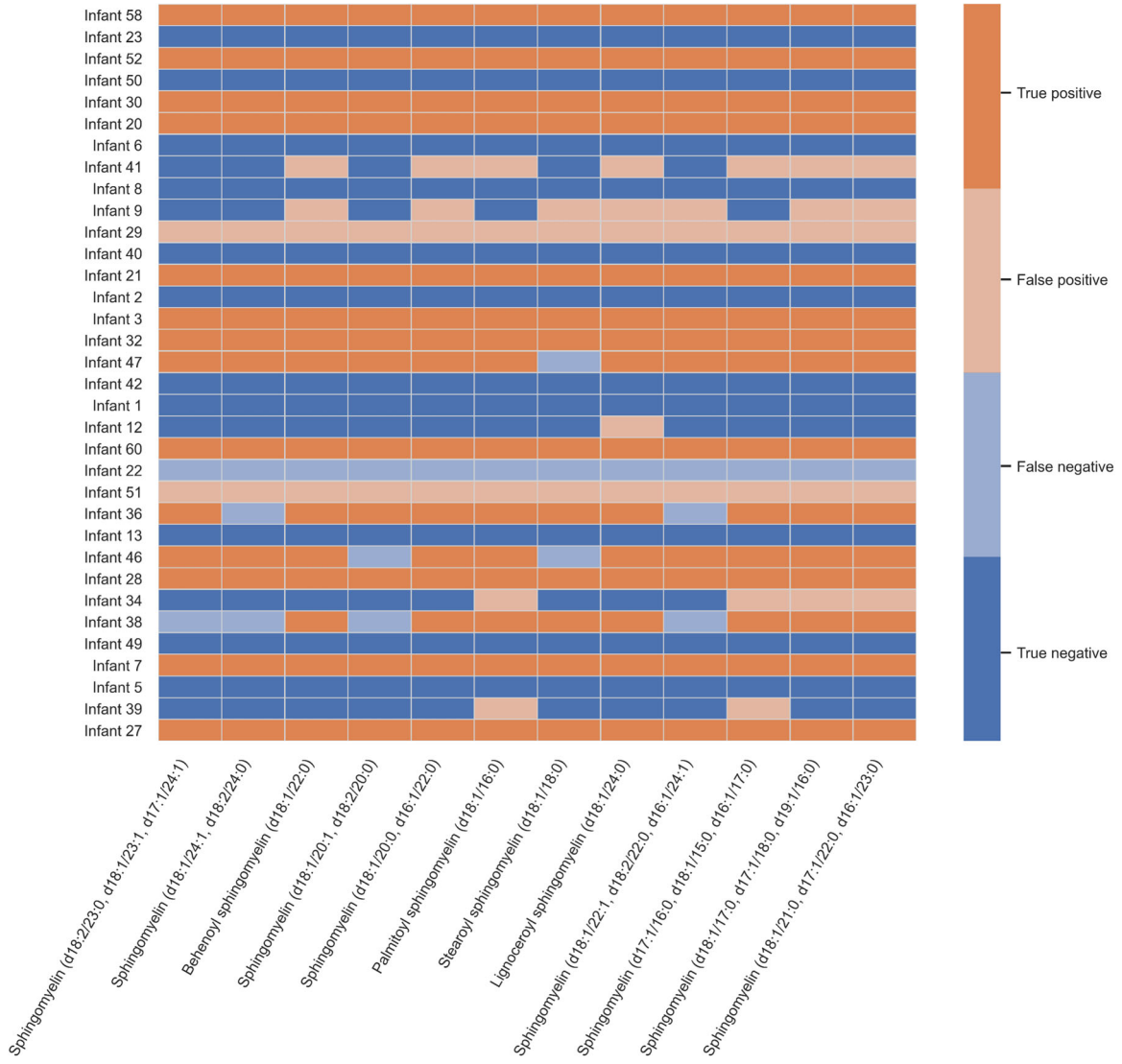

FIGURE 5 The 12 best biomarkers show high agreement across our cohort. There is one infant in particular who contributes the majority of false-negative classifications across all 12 metabolites. Among the infants, we see only false-positive classifications when using an ensemble majority vote across the 12 metabolites. There were four false-positive classifications based on majority vote and one false-negative classification, resulting in an overall accuracy of 85%.

Our results demonstrate that clinical variables recorded at birth, together with frequent testing for biomarkers in the stool, would provide a method for identifying which infants are at risk of developing PNAC. We have proposed a set of simple diagnostic criteria for classifying infants based on their expected risk level for developing PNAC. Early identification of PNAC before elevated conjugated bilirubin levels in the blood would allow NICU medical teams to take early action to limit the occurrence of liver damage in this vulnerable population. The most likely course of action would involve switching an infant at risk of PNAC to a hepatoprotective lipid emulsion such as Omegaven. This precise diagnostic plan would allow medical teams to proactively optimize for infant health outcomes while also helping the NICU to account for other competing objectives, such as cost. Additionally, frequent monitoring of the stool may enable clinicians to confidently optimize energy nutrition with PN for infants at low risk of developing PNAC, which is a known enhancer of health outcomes in the NICU.

Sphingomyelin in the stool was the most predictive fecal metabolomic signature for the risk of PNAC. This lipid plays a role in inflammatory signaling in the GI tract, tight junction maintenance, and the metabolism of nutrients present in the GI tract.[28,29] Additionally, there are several known connections with GI diseases such as necrotizing enterocolitis, ulcerative colitis, and GI hyperpermeability.[28,30] Finally, there are also connections to the GI microbiota. Sphingomyelin has been shown to play a role in the physiology of how probiotics interact with the intestinal lining.[31] The underlying mechanism of why sphingomyelin is predictive of PNAC requires further elucidation.

The diagnostic potential that resides in biological samples that are currently treated as waste in the NICU is immense. Our results demonstrate that stool samples contain measurable biomarkers that are predictive of disease. In the NICU, there is a constant need for more information to help treat and take care of premature infants. Stool and urine represent two additional sources of valuable information that have previously been out of reach because of the complexity of identifying effective biomarkers for disease. However, with the advent of advanced metabolomics and systems biology, there is a new opportunity to advance diagnostic procedures in the NICU past blood tests and monitoring of vitals. PNAC is only one of many devastating diseases in the NICU that may be mitigated by early identification of infants at risk through the use of stool samples.

To our knowledge, this is the first prospective study to investigate metabolite predictors of the risk of developing PNAC. The study was limited by the size and single-center scope of our cohort. Further, our prediction models require validation in an independent cohort and also possibly in a cohort at greater risk of developing PNAC, for example, lower overall gestational age and birth weight. However, accurate classification within our current cohort, together with high correlation of fecal sphingomyelins with serum direct bilirubin levels over time, provides support that these biomarkers hold potential for prediction of infants at risk of developing PNAC.

AUTHOR CONTRIBUTIONS

Thomas J. Moutinho Jr, Jason A. Papin, Sean R. Moore, and Suchitra K. Hourigan designed research; Thomas J. Moutinho Jr, Shira Levy, Deborah A. Powers, Robin L. Baker, Isabel Hefner, Masouma Mohamed, Alaa Abdelghani, and Rajiv Baveja conducted research; Thomas J. Moutinho Jr, Gabriel F. Hanson, and Deborah A. Powers analyzed data; and Thomas J. Moutinho Jr, Shira Levy, Deborah A. Powers, Gabriel F. Hanson, Jason A. Papin, Sean R. Moore, and Suchitra K. Hourigan wrote the article. Thomas J. Moutinho Jr, Jason A. Papin, Sean R. Moore, and Suchitra K. Hourigan had primary responsibility for final content. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

None declared.

FUNDING INFORMATION

Research reported in this publication was supported in part by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR003015 (Moore and Hourigan), the National Institute of Child Health and Human Development under Award Number K23HD099240 (Hourigan), and by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases (Hourigan) of the National Institutes of Health. Additional support was provided by the University of Virginia-Coulter Translational Partnership (Papin and Moore). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DATA AVAILABILITY STATEMENT

Data described in the article, code book, and analytic code will be made available on request pending application and approval.

ORCID

Thomas J. Moutinho http://orcid.org/0000-0002-8940-4970
Sean R. Moore http://orcid.org/0000-0003-1150-6632

REFERENCES

1. Khalaf RT, Sokol RJ. New insights into intestinal failure-associated liver disease in children. *Hepatology*. 2020;71(4):1486-1498.
2. Venigalla S, Gourley GR. Neonatal cholestasis. *Semin Perinatol*. 2004;28(5):348-355.
3. Hsieh M-H, Pai W, Tseng H-I, Yang S-N, Lu C-C, Chen H-L. Parenteral nutrition-associated cholestasis in premature babies: risk factors and predictors. *Pediatr Neonatol*. 2009;50(5):202-207.
4. Alkharfy TM, Ba-Abbad R, Hadi A, Sobaih BH, AlFaleh KM. Total parenteral nutrition-associated cholestasis and risk factors in preterm infants. *Saudi J Gastroenterol*. 2014;20(5):293-296.
5. Christensen RD, Henry E, Wiedmeier SE, Burnett J, Lambert DK. Identifying patients, on the first day of life, at high-risk of developing parenteral nutrition-associated liver disease. *J Perinatol*. 2007;27(5):284-290.
6. Mangalat N, Bell C, Graves A, Imseis EM. Natural history of conjugated bilirubin trajectory in neonates following parenteral nutrition cessation. *BMC Pediatr*. 2014;14(1):298.
7. Park HW, Lee NM, Kim JH, Kim KS, Kim S-N. Parenteral fish oil-containing lipid emulsions may reverse parenteral nutrition-associated cholestasis in neonates: a systematic review and meta-analysis. *J Nutr*. 2015;145(2):277-283.
8. Kasirer Y, Bin-Nun A, Raveh A, Schorrs I, Mimouni FB, Hammerman C. SMOFlipid protects preterm neonates against perinatal nutrition-associated cholestasis. *Am J Perinatol*. 2019;36(13):1382-1386.
9. Casson C, Nguyen V, Nayak P, et al. A comparison of Smoflipid® and Intralipid® in the early management of infants with intestinal failure. *J Pediatr Surg*. 2020;55(1):153-157.
10. Zhgun ES, Ilina EN. Fecal metabolites as non-invasive biomarkers of gut diseases. *Acta Naturae*. 2020;12(2):4-14.
11. Bouatra S, Aziat F, Mandal R, et al. The human urine metabolome. *PLoS One*. 2013;8(9):e73076.
12. Cahova M, Bratova M, Wohl P. Parenteral nutrition-associated liver disease: the role of the gut microbiota. *Nutrients*. 2017;9(9):987.
13. Lee WS, Sokol RJ. Intestinal microbiota, lipids, and the pathogenesis of intestinal failure-associated liver disease. *J Pediatr*. 2015;167(3):519-526.
14. Hourigan SK, Moutinho TJ Jr, Berenz A, et al. Gram-negative microbiota blooms in premature twins discordant for parenteral nutrition associated cholestasis. *J Pediatr Gastroenterol Nutr*. 2020;70(5):640-644.
15. Li Y, Tang R, Leung PSC, Gershwin ME, Ma X. Bile acids and intestinal microbiota in autoimmune cholestatic liver diseases. *Autoimmun Rev*. 2017;16(9):885-896.
16. Fawaz R, Baumann U, Ekong U, et al. Guideline for the evaluation of cholestatic jaundice in infants: joint recommendations of the North American society for pediatric gastroenterology, hepatology, and nutrition and the European society for pediatric gastroenterology, hepatology, and nutrition. *J Pediatr Gastroenterol Nutr*. 2017;64(1):154-168.
17. Almonacid DE, Kraal L, Ossandon FJ, et al. 16S rRNA gene sequencing and healthy reference ranges for 28 clinically relevant microbial taxa from the human gut microbiome. *PLoS One*. 2017;12(5):e0176555.

18. Ramamoorthy S, Levy S, Mohamed M, et al. An ambient-temperature storage and stabilization device performs comparably to flash-frozen collection for stool metabolomics in infants. *BMC Microbiol*. 2021;21(1):59.
19. Loftfield E, Vogtmann E, Sampson JN, et al. Comparison of collection methods for fecal samples for discovery metabolomics in epidemiologic studies. *Cancer Epidemiol Biomarkers Prev*. 2016;25(11):1483-1490.
20. Zierer J, Jackson MA, Kastenmüller G, et al. The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet*. 2018;50(6):790-795.
21. Evans AM, Bridgewater B, Liu Q, et al. High resolution mass spectrometry improves data quantity and quality as compared to unit mass resolution mass spectrometry in high-throughput profiling metabolomics. *Metabolomics*. 2014;4(2):1.
22. Ford L, Kennedy AD, Goodman KD, et al. Precision of a clinical metabolomics profiling platform for use in the identification of inborn errors of metabolism. *J Appl Lab Med*. 2020;5(2):342-356.
23. Evans AM, DeHaven CD, Barrett T, Mitchell M, Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal Chem*. 2009;81(16):6656-6667.
24. DeHaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J Cheminform*. 2010;2(1):1-12.
25. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Meth*. 2016;13(7):581-583.
26. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. 2017;11(12):2639-2643.
27. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37(8):852-857.
28. Bock J, Liebisch G, Schweimer J, Schmitz G, Rogler G. Exogenous sphingomyelinase causes impaired intestinal epithelial barrier function. *World J Gastroenterol*. 2007;13(39):5217-5225.
29. Qi Y, Jiang C, Cheng J, et al. Bile acid signaling in lipid metabolism: metabolomic and lipidomic analysis of lipid and bile acid markers linked to anti-obesity and anti-diabetes in mice. *Biochim Biophys Acta*. 2015;1851(1):19-29.
30. Kurek K, Łukaszuk B, Piotrowska DM, Wiesiołek P, Chabowska AM, Żendzian-Piotrowska M. Metabolism, physiological role, and clinical implications of sphingolipids in gastrointestinal tract. *BioMed Res Int*. 2013;2013(20):e908907.
31. Soo I, Madsen KL, Tejpar Q, et al. VSL#3 probiotic upregulates intestinal mucosal alkaline sphingomyelinase and reduces inflammation. *Can J Gastroenterol*. 2008;22(3):237-242.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Moutinho TJ, Powers DA, Hanson GF, et al. Fecal sphingolipids predict parenteral nutrition–associated cholestasis in the neonatal intensive care unit. *J Parenter Enteral Nutr*. 2022;46:1903-1913. doi:10.1002/jpen.2374