Clinical Infectious Diseases

MAJOR ARTICLE

Late-onset Sepsis in Preterm Infants Can Be Detected Preclinically by Fecal Volatile Organic Compound Analysis: A Prospective, Multicenter Cohort Study

Daniel J. C. Berkhout,[1,2] Britt J. van Keulen,[1] Hendrik J. Niemarkt,[3] Jet R. Bessem,[2] Willem P. de Boode,[4] Veerle Cossey,[5] Neil Hoogenes,[2] Christiaan V. Hulzebos,[6] Ellen Klaver,[2] Peter Andriessen,[3] Anton H. van Kaam,[7,8] Boris W. Kramer,[9] Richard A. van Lingen,[10] Aaron Schouten,[2] Johannes B. van Goudoever,[11,12] Daniel C. Vijlbrief,[13] Mirjam M. van Weissenbruch,[7] Alfian N. Wicaksono[14] James A. Covington,[14] Marc A. Benninga,[1] Nanne K. H. de Boer,[15] and Tim G. J. de Meij[2]

[1]Department of Pediatric Gastroenterology, Emma Children's Hospital/Academic Medical Center, and [2]Department of Pediatric Gastroenterology, VU University Medical Center, Amsterdam, [3]Neonatal Intensive Care Unit, Máxima Medical Center, Veldhoven, and [4]Neonatal Intensive Care Unit, Radboud University Medical Center, Radboud Institute for Health Sciences, Amalia Children's Hospital, Nijmegen, The Netherlands; [5]Neonatal Intensive Care Unit, University Hospitals Leuven, Belgium; [6]Neonatal Intensive Care Unit, Beatrix Children's Hospital/University Medical Center Groningen, [7]Neonatal Intensive Care Unit, VU University Medical Center, and [8]Neonatal Intensive Care Unit, Emma Children's Hospital/Academic Medical Center, Amsterdam, [9]Department of Pediatrics, Maastricht University Medical Center, [10]Neonatal Intensive Care Unit, Amalia Children's Centre/Isala, Zwolle, [11]Department of Pediatrics, Emma Children's Hospital/Academic Medical Center, and [12]Department of Pediatrics, VU University Medical Center, Amsterdam, and [13]Neonatal Intensive Care Unit, Wilhelmina Children's Hospital/University Medical Center Utrecht, The Netherlands; [14]School of Engineering, University of Warwick, United Kingdom; and [15]Department of Gastroenterology and Hepatology, VU University Medical Center, Amsterdam, The Netherlands

Background. The intestinal microbiota has increasingly been considered to play a role in the etiology of late-onset sepsis (LOS). We hypothesize that early alterations in fecal volatile organic compounds (VOCs), reflecting intestinal microbiota composition and function, allow for discrimination between infants developing LOS and controls in a preclinical stage.

Methods. In 9 neonatal intensive care units in the Netherlands and Belgium, fecal samples of preterm infants born at a gestational age ≤30 weeks were collected daily, up to the postnatal age of 28 days. Fecal VOC were measured by high-field asymmetric waveform ion mobility spectrometry (FAIMS). VOC profiles of LOS infants, up to 3 days prior to clinical LOS onset, were compared with profiles from matched controls.

Results. In total, 843 preterm born infants (gestational age ≤30 weeks) were included. From 127 LOS cases and 127 matched controls, fecal samples were analyzed by means of FAIMS. Fecal VOCs allowed for preclinical discrimination between LOS and control infants. Focusing on individual pathogens, fecal VOCs differed significantly between LOS cases and controls at all predefined time points. Highest accuracy rates were obtained for sepsis caused by *Escherichia coli*, followed by sepsis caused by *Staphylococcus aureus* and *Staphylococcus epidermidis*.

Conclusions. Fecal VOC analysis allowed for preclinical discrimination between infants developing LOS and matched controls. Early detection of LOS may provide clinicians a window of opportunity for timely initiation of individualized therapeutic strategies aimed at prevention of sepsis, possibly improving LOS-related morbidity and mortality.

Keywords. neonatology; volatile organic compound; late-onset sepsis; high-field asymmetric waveform ion mobility spectrometry; electronic nose.

Despite a decrease in the incidence of late-onset sepsis (LOS) over the past decade, still 34% of all extremely low birth weight and preterm infants, born at a gestational age <29 weeks, experience at least 1 episode of LOS [1]. Unfortunately, this decline in incidence is not accompanied with a decrease in either LOS associated mortality (12%) or length of hospital stay [1].

Although extensively studied, the exact biological mechanisms of LOS are still poorly understood. Next to indwelling devices, the intestinal microbiota has increasingly been considered to play a pivotal role in LOS etiology. Bacterial dysbiosis, combined with an immature intestinal epithelial barrier and naive immune system, may increase the risk of transmucosal bacterial translocation, ultimately leading to "gut-derived sepsis." Several microbiota studies have demonstrated microbial alterations preceding the clinical onset of LOS, as compared to controls [2–9]. However, a key feature of these studies is the lack of a consistent LOS-specific microbial signature. Furthermore, implementation of intestinal microbiota analysis in daily clinical practice as an early diagnostic biomarker for LOS is not feasible because of high costs and the complexity to generate and interpret microbiota results in a clinically acceptable fashion.

Fecal volatile organic compound (VOC) analysis may serve as alternative test to monitor changes in microbiota and its metabolic activity. VOCs are carbon-based volatiles, which are

Received 26 March 2018; editorial decision 24 April 2018; accepted 27 April 2018; published online June 21, 2018.

Correspondence: D. J. C. Berkhout, Department of Pediatrics, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands (d.berkhout@vumc.nl).

Clinical Infectious Diseases® 2018;XX(XX):1–8

DOI: 10.1093/cid/ciy383

considered to reflect microbial composition and function, as they provide insight into the interaction between gut microbiota and host [10, 11]. Fecal VOCs have previously been described to hold potential as noninvasive diagnostic biomarker of diseases that are associated with alterations in intestinal microbiota composition [12, 13]. Recently, in a small proof of principle study, we have shown that fecal VOCs allow for differentiation between preterm infants with and without LOS, up to 3 days prior to clinical onset of LOS [14]. Because all bacterial species exhibit a unique, metabolic signature [10] and provoke species-specific host-pathogen interactions, fecal VOCs may hypothetically allow for identification of LOS-provoking pathogens at species level.

The primary aim of this study was to investigate whether fecal VOC analysis could discriminate preterm infants developing LOS from controls before the onset of symptoms of LOS, in a large multicenter cohort. The secondary aim was to evaluate whether the discriminative accuracy would increase when focusing on specific causative LOS pathogens.

METHODS

Subjects

For a more detailed description, we would like to refer you to the Supplementay Material section. In short, between October 2014 and January 2017, infants born at a gestational age ≤30 weeks were consecutively included in this prospective cohort study performed at 9 neonatal intensive care units (NICUs) located in the Netherlands and Belgium. The study was approved by all local Institutional Review Boards (protocol A2016.363), and written parental informed consent was obtained from all included patients.

Patient Cohort and Sample Selection

Included preterm infants were identified as LOS cases if all 3 Vermont Oxford criteria for LOS were met [15]. In case a coagulase-negative *Staphylococcus* (CoNS) bacteria was isolated from blood culture, infants were allocated to the LOS group only if C-reactive protein level was at least once ≥10 mg/L within 1 week after clinical onset of LOS, in order to minimize the risk of including infants with contaminated blood cultures [16]. Fecal samples obtained at 3 (t_{-3}), 2 (t_{-2}), and 1 (t_{-1}) day(s) prior to clinical LOS onset were used for further VOC analysis.

Each LOS case was strictly matched to one non-LOS control based on the following criteria: center of birth, gestational age, postnatal age, birth weight, days exposed to antibiotics prior to t_0, and enteral feeding type (breastmilk vs formula feeding).

VOC Comparisons

Fecal VOC profiles from cases and controls were compared at each predefined time point in days (t_{-3}, t_{-2} and t_{-1}) and, as extremely preterm infants may not pass stool daily, in an analysis including the last obtained sample prior to t_0. Multiple VOC profile comparisons were performed at these measurement points:

1. In the first analysis, the most frequently cultured species of Gram-positive (except CoNS; *Staphylococcus aureus*), Gram-negative (*Escherichia coli*), and CoNS (*Staphylococcus epidermidis*) were identified. Fecal samples from cases and controls were compared at the predefined time points and in an analysis including only the last sample obtained prior to LOS onset. In case both CoNS and a non-CoNS species were obtained from one single blood culture, the LOS case was allocated to the non-CoNS group [15]. Cultures containing 2 different CoNS species or multiple non-CoNS species in one single blood culture were excluded from further VOC analysis.
2. For the second analysis, pathogens were allocated to either the Gram-positive (except CoNS), Gram-negative or CoNS group, comparing VOC profiles from the last obtained sample prior to t_0 with their corresponding matched control sample. Here, if both Gram-positive and Gram-negative bacteria were obtained from the blood culture, cases were included in both Gram-negative and Gram-positive analysis. Cultures containing CoNS and non-CoNS pathogens were excluded from the CoNS analysis, because CoNS is often considered to be a contaminant in these cases [15].
3. The third analysis included VOC profiles from all LOS pathogens combined and were compared with their matched controls to assess whether the overall VOC signal in LOS differed from controls. Both cross-sectional analysis per time interval (t_{-3}, t_{-2}, and t_{-1}) and an analysis including only the VOC profile of the last fecal sample produced prior to LOS onset was performed.
4. Fourth, to assess the potential influence of center of birth on fecal VOC outcome, we performed 3 further analyses. Each analysis included only those VOC profiles of the last fecal sample obtained prior to LOS from infants born at one of the 3 centers who included the highest rates of LOS cases (centers 1, 2, and 3).
5. Because blood cultures containing more than one different pathogen are often considered to be contaminated, an additional analysis was performed including only cases with monocultures and compared them with their matched controls.

VOC Analysis by FAIMS Technique

Fecal samples were analyzed using the FAIMS technique (Lonestar®, Owlstone, Cambridge, England) and for a detailed description of the FAIMS technique, we refer to previous studies and the Supplementay Material [13, 17].

Statistical Analysis

Demographic and clinical variables were compared using a χ^2, independent *t*-test, or nonparametric test where appropriate. A *P*-value of <.05 was considered statistically significant. For the VOC-profile analyses, after reduction of the raw data

using 2D discrete wavelet transform (a common form of data compression), a Wilcoxon rank-sum test was used to calculate *P*-values [18]. Subsequently, 4 classification algorithms, Random Forest, Sparse Logistic Regression, Support Vector Machine, and Gaussian Process Classifier, were used to produce class prediction inside a 10-fold cross validation with 90% of the data as a training set and the remaining 10% as a test set (Supplementay Material) [19]. Receiver operating characteristic (ROC) curves with corresponding area under the curves (AUCs), sensitivity, specificity, positive predictive value and negative predictive value were produced for each model, selecting the most significant model for notation.

RESULTS

A total of 843 preterm infants were enrolled in this study, of whom 127 both met the criteria for LOS and had sufficient fecal material available for VOC analysis (Figure 1). Patient characteristics and cultured pathogens from the infants excluded based on unavailability of fecal material were similar to the characteristics and pathogens of the current study. In the present study, the sepsis and control groups did not differ with respect to demographic and clinical variables (Table 1). Variation in initial empirical antibiotic treatment between centers is depicted in Supplementary Table 1. In 14/127 (11%) cases, more than one different pathogen was isolated from the blood culture. Focusing on monocultures, *S. epidermidis* was the most frequently isolated pathogen (n = 42), followed by *Staphylococcus capitis* (n = 17), *S. aureus* (n = 19), and *E. coli* (n = 11) (Table 2).

Fecal Volatile Organic Compound Analysis

An overview of the fecal VOC outcomes is depicted in Table 3, Supplementary Table 2 and Supplementary Figure 1. The median date of birth with corresponding interquartile range (IQR) was 5 May 2016 (427 days) for cases and 17 May 2016 (311 days) for controls and could be considered as a reflection of the median sample storage time per study group. Summarized, we observed that fecal VOC profiles of infants with LOS caused by *S. aureus* and *E. coli*, differed significantly from matched controls infants at all 3 predefined time points. Corresponding AUCs in the analysis including only *S. aureus* were .85, .70, .80, and in the analysis including *E. coli* .88, .99, .86 *at* t_{-3}, $t_{-2,}$ and t_{-1} respectively. Pooling the samples obtained most adjacent to t_0 from all Gram-negative (AUC = .77) or Gram-positive pathogens allowed for discrimination between cases and controls. In the subgroup analysis including VOC profiles of all CoNS together, and the analysis focusing on the last obtained sample prior onset from the most frequently cultured species within this CoNS subgroup (*S. epidermidis*), fecal VOCs did not allow for discrimination between LOS cases and controls. In contrast, separating these samples based on the time points they were obtained, fecal VOCs allowed for discrimination between cases and controls with corresponding AUC of .90, .78, .63 at t_{-3}, $t_{-2,}$ and $t_{-1,}$ respectively.

If all LOS infants were included, irrespective of the cultured pathogen, fecal VOC analysis only allowed for a statistically significant discrimination at t_{-1}. If only the last produced fecal sample prior t_0 was included in the analysis, fecal VOC analysis remained statistically significantly different.

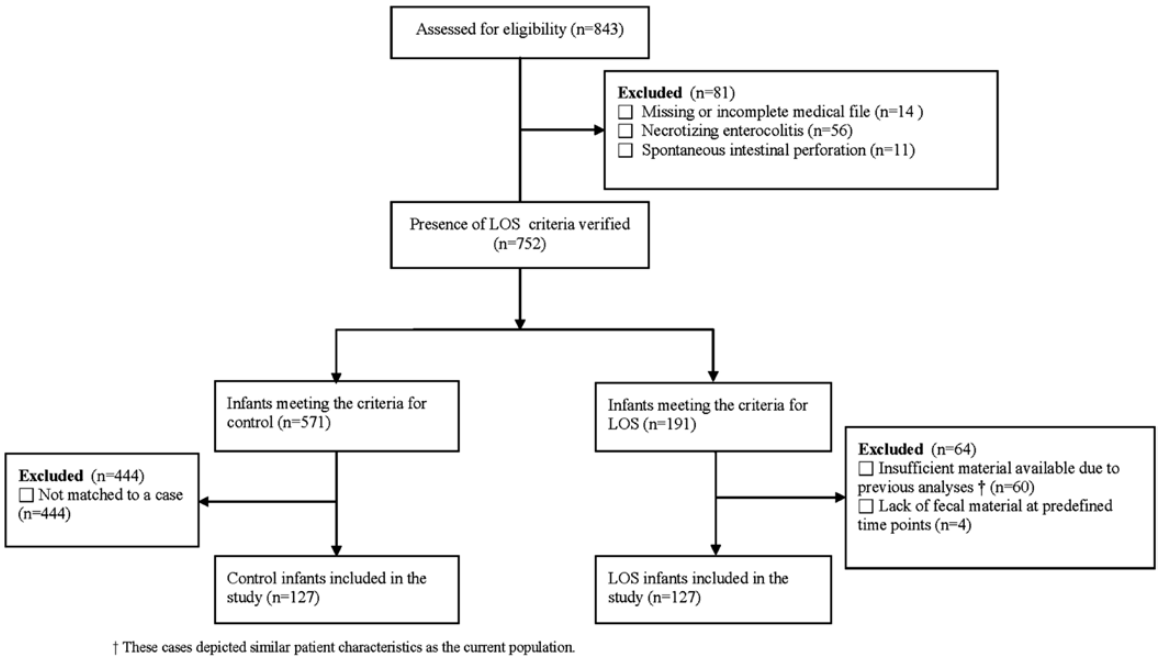

Figure 1. Flow-chart of participants in the study. Abbreviation: LOS, late-onset sepsis.

Table 1. Subject Characteristics of Cases and Controls

	Sepsis (n = 127)	Controls (n = 127)	*P* Value
Gestational age (median [IQR]), weeks + days [days]	26 + 6 [19]	27 + 0 [14]	.384
Birth weight (median [IQR]), g	920 [365]	964 [280]	.280
Sex			
Male (n [%])	66 [52.0]	70 [55.1]	.615
Delivery mode			
Vaginal delivery (n [%])	67 [52.8]	62 [48.8]	.530
Multiple births (n [%])	57 [44.9]	48 [37.8]	.251
Postnatal age at t_0 (median [IQR]), days	9 [7]	n.a.	n.a.
Antibiotic exposure prior t_0 (n [%])	118 [92.9]	119 [93.7]	.802
Antibiotic days (median [IQR])	4 [3]	4 [3]	.911
Enteral feeding type prior t_0[a] (n [%])			.857
Breastmilk fed	53 [45.7]	57 [49.1]	
Formula fed	31 [26.7]	30 [25.9]	
Combination	32 [27.6]	29 [25.0]	
Mortality (n [%])	8 [6.3]	2 [1.6]	.053
Age of death (median [IQR]), days	15.5 [7.8]	9.5	.711

[a] Variables were not retrievable from the medical records in one participating center (n = 22).

Abbreviation: IQR, interquartile range.

Three additional analyses were performed including infants born at one of the 3 centers, which included the largest number of LOS cases (n = 39, n = 25, n = 16, respectively). Fecal VOCs allowed for statistically significant discrimination in center 1 (AUC [±95% confidence interval (CI)], *P*-value, sensitivity, specificity;.76 [.63–.90], .0005, .76, .72) and center 2 (.86 [.71–1], .0001, .88, .88) but not in center 3 (.62 [.50–.75], .9693, .74, .59). At least one CoNS was isolated in 33 of 39 (85%) sepsis cases born in center 3, whereas CoNS was identified in 13/25 (52%) cases and 11/16 (69%) cases in, respectively, center 1 and center 2.

DISCUSSION

In this prospective multicenter cohort study, we investigated the diagnostic potential of fecal VOCs in the preclinical detection of LOS. In general, VOC profiles of 127 infants with LOS could be discriminated from controls before the clinical diagnosis was established. Specifically, LOS caused by *E. coli, S. aureus*, and *S. epidermidis* could be differentiated from their matched controls with high predictive value, up to 3 days before the clinical onset of LOS.

Previously, we have demonstrated that fecal VOCs allowed for differentiation between infants developing LOS and matched controls with AUCs of 70.2, 77.7, and 70.4 at, respectively, 3, 2, and 1 day prior to LOS onset but not earlier [14]. However, that study was limited by a wide variety of pathogens and a small sample size (n = 36 cases). In addition, VOC analyses in that study were performed by a Cyranose320® eNose (Cyranose320®, Sensigent, US) harboring 32 unique conducting sensors, whereas the currently applied FAIMS technique allows for over 52.000 individual data points [18, 20], based on the mobility of ionized molecules in an electrical field.

In the present study, including a larger number of LOS cases, discriminative accuracies improved compared to the results obtained in the previous study, in case of *E. coli* (n = 14) and *S. aureus* (n = 21) LOS. This suggests the presence of a unique, species-specific fecal VOC profile in LOS. It has been demonstrated, mainly using in vitro studies, that individual bacterial species are characterized by production of a unique VOC fingerprint [10, 21]. We therefore hypothesized that changes in the composition of the microbiota prior to LOS onset may have largely contributed to the detected alterations in fecal VOC profiles. To date, the intestinal microbiota has increasingly been recognized as an important factor in LOS etiology in preterm infants. Interestingly, an increased abundance of the LOS causing pathogen preceding onset, including *E. coli* [2, 7, 8] and *S. aureus* [5, 7], has been demonstrated in several studies.

By pooling both Gram-negative pathogens (n = 27) and Gram-positive pathogens (n = 28), discriminative accuracy slightly decreased compared to the analysis involving unique species. This finding may be explained by the observation that Gram-positive and Gram-negative pathogens depict specific metabolic processes, allowing for discrimination between both subgroups by VOC analysis [22].

As expected, including fecal samples obtained most adjacent to t_0 from all included LOS infants (n = 127) in the analysis, discriminative accuracy decreased even further. This may be explained by the wide variety of different LOS pathogens in this study, each pathogen exhibiting a distinct VOC profile. Presumably, the presence of discriminative VOCs in septic infants resulting from production of local and systemic inflammatory biomarkers allowed for the discrimination between LOS and control infants. Another possible explanation is that absence of VOCs originating from commensals characterizing healthy state in controls, including anaerobic bacteria of the

Table 2. Isolated Pathogens (n [%]) From Blood Cultures in 127 Sepsis Patients

	n [%]
Monomicrobial cultures	113 [89]
CoNS	
- *Staphylococcus epidermidis*	42 [33.1]
- *Staphylococcus capitis*	17 [13.4]
- *Staphylococcus haemolyticus*	4 [3.1]
- *Staphylococcus warneri*	2 [1.6]
- *Staphylococcus hominis*	1 [0.8]
- Coagulase negative *Staphylococcus*	1 [0.8]
Gram-positive pathogens[a]	
- *Staphylococcus aureus*	19 [15.0]
- *Bacillus cereus*	2 [1.6]
- *Enterococcus faecalis*	1 [0.8]
- Group B *Streptococcus*	1 [0.8]
- Group C *Streptococcus*	1 [0.8]
Gram-negative pathogens	
- *Escherichia coli*	11 [8.7]
- *Enterobacter cloacae*	4 [3.1]
- *Serratia marcescens*	3 [2.4]
- *Enterobacter aerogenes*	1 [0.8]
- *Klebsiella pneumoniae*	1 [0.8]
- *Serratia liquefaciens*	1 [0.8]
Fungal pathogens	
- *Candida Albicans*	1 [0.8]
Cultures with ≧2 different pathogens	14 [11.0]
Gram-negative pathogens + Gram -positive pathogens[a] + CoNS	
- *Klebsiella ornithinolytica* + *Enterococcus faecalis* + *Staphylococcus epidermidis*	1 [0.8]
Gram-negative pathogens + Gram-positive pathogens[a]	
- *Klebsiella pneumoniae* + *Staphylococcus aureus*	1 [0.8]
Gram-negative pathogens + CoNS	
- *Escherichia coli* + *Staphylococcus epidermidis*	3 [2.4]
- *Enterobacter cloacae* + *Staphylococcus epidermidis*	1 [0.8]
Gram-positive pathogens[a] + CoNS	
- *Staphylococcus aureus* + *Staphylococcus epidermidis*	2 [1.6]
CoNS + CoNS	
- *Staphylococcus capitis* + *Staphylococcus epidermidis*	3 [2.4]
- *Staphylococcus epidermidis* + *Staphylococcus haemolyticus*	2 [1.6]
- *Staphylococcus capitis* + *Staphylococcus haemolyticus* + *Staphylococcus warneri*	1 [0.8]

Abbreviations: CoNS, Coagulase negative *Staphylococcus*.

[a]Not including coagulase negative *Staphylococcus*

genus *Bifidobacterium* [4, 7], allowed for the discrimination in this specific analysis.

In the current study, focusing on the last sample obtained prior to disease onset, fecal VOCs did not allow for discrimination between sepsis caused by either *S. epidermidis* (n = 42) or CoNS (n = 73), in general, and strictly matched controls. Because CoNS species are considered predominant components of both dermal [23] and intestinal microbiome [24], we hypothesize that, in at least a part of the cases, not intestinal mucosal translocation but skin invasive procedures may have been the source of CoNS bacteremia. Consequently, bacterial dysbiosis in the intestines preceding a LOS episode may not have occurred, hampering discrimination between cases and controls based on fecal VOCs in this particular group. This hypothesis is underlined by a recent study, demonstrating the gastrointestinal tract to be dominated by the LOS causing pathogen prior to onset, except for *Staphylococcus epidermis* [7]. Interestingly, separating these samples based on the time points they were obtained, fecal VOCs allowed for discrimination between cases with *S. epidermidis* sepsis and controls at each individual time-point. It has previously been described that the center of birth has a significant influence on fecal VOC profiles, possibly resulting from center-specific protocols on feeding patterns and choice of antibiotics [25]. This finding was confirmed in the present study, because focusing on VOC profiles of infants from the same center of birth resulted in an increased discriminative accuracy between LOS and controls, the only exception being center 3. We hypothesized that this apparent discrepancy resulted from the presence of CoNS in the vast majority of the sepsis cases (85%) in that particular center, whereas CoNS species were less frequently isolated from blood culture in the remaining two centers. As described previously, detection of CoNS by fecal VOC analysis seems to be less feasible than other pathogens, possibly by a different site of entry.

In the current study, observations are highly dependent on the selected statistical analytical supervised method. Because the performance of various prediction models has been shown to vary considerably, it has been questioned whether a certain classifier model may have universal applicability in different populations [19]. Further studies are needed to assess which model provides optimal accuracy in different research and clinical settings.

Strengths of the current study are the large number of included cases in this multicenter, prospective designed study with daily collection of samples and detailed collection on clinical variables, allowing for applying a matching procedure of cases and controls. This study also has several limitations. First, differences in storage time existed between the first collected samples and the samples collected at the end of the study period. This difference in storage time may have caused degradation of volatiles and consequently has influenced VOC outcome. However, we believe that these potential effects on presented results are limited, since this accounted for both cases and controls [26]. Second, microbiota analyses were not performed in the current study. Consequently, study design did not provide prove for our hypothesis that the LOS causing pathogens is abundantly present in the intestines preceding clinical onset. Third, based on our previous study [14], we have only analyzed fecal samples obtained within 3 days prior to LOS onset. Possibly, VOC differences are present already more than 3 days prior to LOS onset, providing a larger window of opportunity to prevent LOS. Fourth, although the eNose instrument allows for bedside and

Table 3. Performance Characteristics for the Discrimination of Late-onset Sepsis and Matched Controls Using Fecal Volatile Organic Compounds

Analysis	Sepsis Samples[a] (n)	*P* Value	AUC (± 95% CI)	Sensitivity (± 95% CI)	Specificity (± 95% CI)	PPV	NPV	Applied Method
Escherichia coli[b]	14	.0002	0.87 (0.74–1)	0.93 (0.66–1)	0.71 (0.42–0.92)	0.76	0.91	Gaussian Process
Escherichia coli t_{-1}	13	.0006	0.86 (0.71–1)	0.92 (0.64–1)	0.77 (0.46–0.95)	0.8	0.91	Support Vector Machine
Escherichia coli t_{-2}	9	<.0001	0.99 (0.95–1.0)	1 (0.66–1)	0.89 (0.52–1.0)	0.9	1	Gaussian Process
Escherichia coli t_{-3}	11	.0013	0.88 (0.72–1)	0.91 (0.59–1.0)	0.82 (0.48–0.98)	0.83	0.9	Random Forest
S. aureus[b]	21	.0191	0.69 (0.52–0.85)	0.76 (0.53–0.92)	0.62 (0.38–0.82)	0.67	0.72	Gaussian Process
S. aureus t_{-1}	15	.0016	0.8 (0.64–0.96)	0.73 (0.45–0.92)	0.8 (0.52–0.96)	0.79	0.75	Support Vector Machine
S. aureus t_{-2}	13	.0406	0.7 (0.5–0.91)	0.85 (0.55–0.98)	0.62 (0.32–0.86)	0.69	0.8	Random Forest
S. aureus t_{-3}	16	.0002	0.85 (0.7–1)	0.88 (0.62–0.98)	0.81 (0.54–0.96)	0.82	0.87	Gaussian Process
S. epidermidis[b]	42	.9823	0.63 (0.51–0.75)	0.74 (0.58–0.86)	0.55 (0.39–0.7)	0.62	0.68	Support Vector Machine
S. epidermidis t_{-1}	35	.0308	0.63 (0.5–0.76)	0.54 (0.37–0.71)	0.71 (0.54–0.85)	0.66	0.61	Gaussian Process
S. epidermidis t_{-2}	22	.0006	0.78 (0.64–0.92)	0.82 (0.60–0.95)	0.68 (0.45–0.86)	0.72	0.79	Sparse Logistic Regression
S. epidermidis t_{-3}	19	<.0001	0.90 (0.79–1.0)	0.84 (0.6–0.97)	0.89 (0.67–0.99)	0.89	0.85	Random Forest
Gram-negative bacteria[b]	27	.0030	0.77 (0.63–0.9)	0.78 (0.58–0.91)	0.81 (0.62–0.94)	0.81	0.79	Support Vector Machine
Gram-positive bacteria[b]	28	.0007	0.74 (0.61–0.88)	0.75 (0.55–0.89)	0.75 (0.55–0.89)	.75	.75	Sparse Logistic Regression
CoNS[b]	73	.1077	0.56 (0.47–0.65)	0.56 (0.44–0.68)	0.6 (0.48–0.72)	0.59	0.58	Random Forest
t_{-1} to t_{-3}[b]	127	.0437	0.56 (0.49–0.63)	0.69 (0.6–0.76)	0.44 (0.35–0.53)	0.55	0.58	Random Forest
t_{-1}	105	.0249	0.58 (0.5–0.66)	0.61 (0.51–0.7)	0.55 (0.45–0.65)	0.58	0.59	Random Forest
t_{-2}	78	.9898	0.61 (0.52–0.7)	0.91 (0.82–0.96)	0.29 (0.2–0.41)	0.56	0.77	Gaussian Process
t_{-3}	78	.9791	0.59 (0.51–0.68)	0.55 (0.43–0.66)	0.62 (0.5–0.72)	0.59	0.58	Random Forest
Mono-cultures[b]	113	.0079	0.59 (0.52–0.67)	0.81 (0.72–0.87)	0.4 (0.31–0.49)	0.57	0.67	Random Forest

Corresponding Area Under the Curves, Sensitivity, Specificity, Positive and Negative Predictive Values are Displayed

Abbreviations: AUC ± 95% CI, area under the curve with 95% confidence interval; CoNS, coagulase negative *Staphylococcus*; NPV, negative predictive value; PPV, positive predictive value; S, *Staphylococcus*.

[a]Corresponding number of fecal samples from controls were analyzed.

[b]for this analysis only the last fecal sample produced prior to late-onset sepsis was used.

relative inexpensive VOC analysis of complex gaseous mixtures, they do not allow for the identification of individual VOCs. Identification of individual, LOS-specific volatiles by means of chemical analytical techniques could allow for development of a tailor-made eNose instrument, with increased accuracy compared to the non-primed eNose in the present study. Because the currently applied eNose focuses on the entire spectrum of VOCs and is not specifically designed to detect LOS in a preclinical phase, we hypothesize that including more subjects in the analysis would result in an increase in the number of confounding VOCs, deriving from nonrelevant and nonavoidable environmental and host-specific sources. Potentially, this may have caused the decrease in discriminative accuracy in the analysis comparing the last obtained fecal samples prior to t_0 from infants with a *S. aureus* sepsis, compared to the analysis including the t_{-1} samples. Future studies should focus on identification of discriminative and species-specific VOCs by means of chemical analytical techniques. This may allow for development of LOS-specific eNose sensors to be applied as a diagnostic tool for preclinical detection of LOS in clinical practice, simultaneously providing information about the causative agent. Daily analysis of fecal samples obtained from the diaper would ultimately allow for the early detection and consequently timely intervention of LOS in preterm infants, eventually resulting in a decrease in LOS-related morbidity and mortality. Hypothetically, these sensors could also be incorporated in an incubator, continuously analyzing the VOCs deriving from the infants. In conclusion, we demonstrated that fecal VOC analysis allowed for the preclinical discrimination between infants developing LOS and matched controls, up to three days prior to LOS onset. In particular, the highly pathogenic *E. coli* and *S. aureus* were detectable preclinically with high accuracy. Preclinical detection of LOS may provide clinicians a window of opportunity for timely initiation of individualized therapeutic strategies, for example, narrow spectrum antibiotics, aimed at prevention of sepsis and might ultimately decrease overall morbidity and mortality in preterm born infants.

Supplementary Data

Supplementary materials are available at *Clinical Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Authors' contributions. D. B.: Study concept and design, acquisition of data, performing measurements, interpretation of data; drafting of the manuscript, wrote first draft of manuscript. B. v. K.**: S**tudy concept and design, acquisition of data, critical revision of the manuscript. H. **N.**: Study concept and design, acquisition of data, drafting of the manuscript. J. B.**:** Data acquisition, data interpretation, critical revision of the manuscript. W. d. B.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. V. C.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. N. H.**: D**ata acquisition, data interpretation, critical revision of the manuscript. C. H.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. E. K.**: D**ata acquisition, data interpretation, critical revision of the manuscript. P. A.: Study design, interpretation of data, critical revision of the manuscript. A. v. K.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. B. K.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. R. v. L.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. A. S.**: D**ata acquisition, data interpretation, critical revision of the manuscript. J. v. G.: Study design, data acquisition, data interpretation, critical revision of the manuscript. D. V.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. M. v. W.: Study design, acquisition of data, interpretation of data, critical revision of the manuscript. A. W.: Analysis and interpretation of data, drafting of the manuscript. J. C.: Analysis and interpretation of data, drafting of the manuscript. M. B.: Study design, interpretation of data, critical revision of the manuscript. N. d. B.: Study concept and design, interpretation of data, critical revision of the manuscript. T. d. M.**: S**tudy concept and design, acquisition of data, analysis and interpretation of data; drafting of the manuscript.

In addition, all authors approved the submitted version of the manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Disclaimer. None of the coauthors received an honorarium, grant or other form of payment for the production of this manuscript

Funding. This work was supported by unrestricted grants from the Maag Lever Darm Stichting, Landelijke Vereniging van Crematoria (Dr. C.J. Vaillant Fonds), Zeldzame Ziekte Fonds and Christine Bader Stichting Irene Kinderziekenhuis.

Potential conflicts of interest. J. C. received a grant from IMSPEX Diagnostics. All other authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Greenberg RG, Kandefer S, Do BT, et al. Late-onset sepsis in extremely premature infants: 2000–2011. Pediatr Infect Dis J **2017**; 36:774–9.
2. Carl MA, Ndao IM, Springman AC, et al. Sepsis from the gut: the enteric habitat of bacteria that cause late-onset neonatal bloodstream infections. Clin Infect Dis **2014**; 58:1211–8.
3. Madan JC, Salari RC, Saxena D, et al. Gut microbial colonisation in premature neonates predicts neonatal sepsis. Arch Dis Child Fetal Neonatal Ed **2012**; 97:F456–62.
4. Mai V, Torrazza RM, Ukhanova M, et al. Distortions in development of intestinal microbiota associated with late onset sepsis in preterm infants. PLoS One **2013**; 8:e52876.
5. Shaw AG, Sim K, Randell P, et al. Late-onset bloodstream infection and perturbed maturation of the gastrointestinal microbiota in premature infants. PLoS One **2015**; 10:e0132923.
6. Soeorg H, Huik K, Parm U, et al. Genetic relatedness of coagulase-negative staphylococci from gastrointestinal tract and blood of preterm neonates with late-onset sepsis. Pediatr Infect Dis J **2013**; 32:389–93.
7. Stewart CJ, Embleton ND, Marrs ECL, et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. Microbiome **2017**; 5:75.
8. Taft DH, Ambalavanan N, Schibler KR, et al. Center variation in intestinal microbiota prior to late-onset sepsis in preterm infants. PLoS One **2015**; 10:e0130604.
9. Tarr PI, Warner BB. Gut bacteria and late-onset neonatal bloodstream infections in preterm infants. Semin Fetal Neonatal Med **2016**; 21:388–93.
10. Bos LD, Sterk PJ, Schultz MJ. Volatile metabolites of pathogens: a systematic review. PLoS Pathog **2013**; 9:e1003311.
11. Raman M, Ahmed I, Gillevet PM, et al. Fecal microbiome and volatile organic compound metabolome in obese humans with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol **2013**; 11:868–75 e1-3.
12. de Meij TG, van der Schee MP, Berkhout DJ, et al. Early detection of necrotizing enterocolitis by fecal volatile organic compounds analysis. J Pediatr **2015**; 167:562–7 e1.
13. van Gaal N, Lakenman R, Covington J, et al. Faecal volatile organic compounds analysis using field asymmetric ion mobility spectrometry: non-invasive diagnostics in paediatric inflammatory bowel disease. J Breath Res **2017**; 12:016006.

14. Berkhout DJ, Niemarkt HJ, Buijck M, et al. Detection of sepsis in preterm infants by fecal volatile organic compounds analysis: a proof of principle study. J Pediatr Gastroenterol Nutr **2016**; 65:e47–52.
15. Manual of operations: part 2 data definitions & infant data forms. October 2015 ed. Vermont Oxford Network, **2016**.
16. Hofer N, Zacharias E, Müller W, Resch B. An update on the use of C-reactive protein in early-onset neonatal sepsis: current insights and new tasks. Neonatology **2012**; 102:25–36.
17. Bomers MK, Menke FP, Savage RS, et al. Rapid, accurate, and on-site detection of *C. difficile* in stool samples. Am J Gastroenterol **2015**; 110:588–94.
18. Covington JA, van der Schee MP, Edge AS, Boyle B, Savage RS, Arasaradnam RP. The application of FAIMS gas analysis in medical diagnostics. Analyst **2015**; 140:6775–81.
19. Ai L, Tian H, Chen Z, Chen H, Xu J, Fang JY. Systematic evaluation of supervised classifiers for fecal microbiota-based prediction of colorectal cancer. Oncotarget **2017**; 8:9546–56.
20. Arasaradnam RP, Ouaret N, Thomas MG, et al. Evaluation of gut bacterial populations using an electronic e-nose and field asymmetric ion mobility spectrometry: further insights into 'fermentonomics'. J Med Eng Technol **2012**; 36:333–7.
21. Palma SICJ, Traguedo AP, Porteira AR, Frias MJ, Gamboa H, Roque ACA. Machine learning for the meta-analyses of microbial pathogens' volatile signatures. Sci Rep **2018**; 8:3360.
22. Dolch ME, Janitza S, Boulesteix AL, et al. Gram-negative and -positive bacteria differentiation in blood culture samples by headspace volatile compound analysis. J Biol Res (Thessalon) **2016**; 23:3.
23. Bialkowska-Hobrzanska H, Jaskot D, Hammerberg O. Molecular characterization of the coagulase-negative staphylococcal surface flora of premature neonates. J Gen Microbiol **1993**; 139:2939–44.
24. Jacquot A, Neveu D, Aujoulat F, et al. Dynamics and clinical evolution of bacterial gut microflora in extremely premature patients. J Pediatr **2011**; 158:390–6.
25. Berkhout DJ, Benninga MA, van Stein RM, et al. Effects of sampling conditions and environmental factors on fecal volatile organic compound analysis by an electronic nose device. Sensors (Basel) **2016**; 16.
26. Forbes SL, Rust L, Trebilcock K, Perrault KA, McGrath LT. Effect of age and storage conditions on the volatile organic compound profile of blood. Forensic Sci Med Pathol **2014**; 10:570–82.