

# Optimized Sampling Conditions for Fecal Volatile Organic Compound Analysis by Means of Field Asymmetric Ion Mobility Spectrometry

Sofie Bosch,<sup>\*,†,‡,§</sup> Sofia el Manouni el Hassani,<sup>‡,§</sup> James A. Covington,<sup>||</sup> Alfian N. Wicaksono,<sup>||</sup> Marije K. Bomers,<sup>§</sup> Marc A. Benninga,<sup>⊥</sup> Chris J. J. Mulder,<sup>†</sup> Nanne K. H. de Boer,<sup>†,⊖</sup> and Tim G. J. de Meij<sup>‡,⊖</sup>

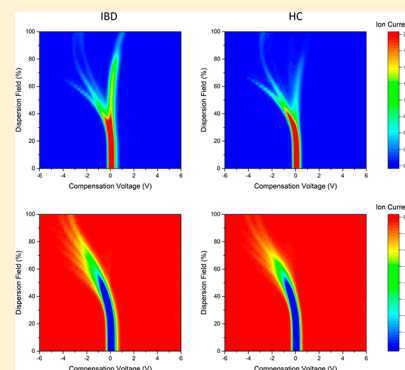
<sup>†</sup>Department of Gastroenterology and Hepatology, <sup>‡</sup>Department of Pediatric Gastroenterology, and <sup>§</sup>Department of Internal Medicine, VU University Medical Center, Amsterdam, The Netherlands

<sup>||</sup>School of Engineering, University of Warwick, Coventry, United Kingdom

<sup>⊥</sup>Department Pediatric Gastroenterology, Emma Children's Hospital/Academic Medical Center, Amsterdam, The Netherlands

## S Supporting Information

**ABSTRACT:** Fecal volatile organic compounds (VOCs) are increasingly considered to be potential noninvasive, diagnostic biomarkers for various gastrointestinal diseases. Knowledge of the influence of sampling conditions on VOC outcomes is limited. We aimed to evaluate the effects of sampling conditions on fecal VOC profiles and to assess under which conditions an optimal diagnostic accuracy in the discrimination between pediatric inflammatory bowel disease (IBD) and controls could be obtained. Fecal samples from de novo treatment-naïve pediatric IBD patients and healthy controls (HC) were used to assess the effects of sampling conditions compared to the standard operating procedure (reference standard), defined as 500 mg of sample mass diluted with 10 mL tap water, using field asymmetric ion mobility spectrometry (FAIMS). A total of 17 IBD (15 CD (Crohn's disease) and 2 UC (ulcerative colitis)) and 25 HC were included. IBD and HC could be discriminated with high accuracy (accuracy = 0.93, AUC = 0.99,  $p < 0.0001$ ). A smaller fecal sample mass resulted in a decreased diagnostic accuracy (300 mg accuracy = 0.77, AUC = 0.69,  $p = 0.02$ ; 100 mg accuracy = 0.70, AUC = 0.74,  $p = 0.003$ ). A loss of diagnostic accuracy was seen toward increased numbers of thaw–freeze cycles (one cycle, accuracy = 0.61, AUC = 0.80,  $p = 0.0004$ ; two cycles, accuracy = 0.64, AUC = 0.56,  $p = 0.753$ ; and three cycles, accuracy = 0.57, AUC = 0.50,  $p = 0.5101$ ) and when samples were kept at room temperature for 180 min prior to analysis (accuracy = 0.60, AUC = 0.51,  $p = 0.46$ ). Diagnostic accuracy of VOC profiles was not significantly influenced by storage duration differences of 20 months. The application of a 500 mg sample mass analyzed after one thaw–freeze cycle showed the best discriminative accuracy for the differentiation of IBD and HC. VOC profiles and diagnostic accuracy were significantly affected by sampling conditions, underlining the need for the implementation of standardized protocols in fecal VOC analysis.



The analysis of volatile organic compounds (VOCs) is a relatively new technique within the field of metabolomics. VOCs are carbon-based chemicals originating from both physiological and pathophysiological processes in the human body. Fecal VOCs are considered to reflect microbiota composition, function, and interaction with the host.<sup>1,2</sup> They are increasingly considered to have potential as a biomarker in the diagnostic workup and monitoring of various gastrointestinal diseases, e.g., inflammatory bowel disease (IBD), colorectal cancer, and even sepsis.<sup>3–11</sup> Various studies have demonstrated the diagnostic potential of VOCs in both pediatric and adult IBD populations by analyzing VOCs deriving from urine, exhaled breath, and feces.<sup>6,12–14</sup> The majority of studies on fecal VOCs has been performed using gas chromatography/mass spectrometry (GC/MS), allowing for the identification of individual VOCs on a molecular level. This technique is

expensive, time consuming, and requires specialized personnel and is therefore not suitable for utilization in a clinical setting.<sup>15</sup> Pattern-recognition-based techniques, such as electronic noses (eNose) and field asymmetric ion mobility spectrometry (FAIMS), are examples of instruments that are less expensive and faster, allowing for their application as a noninvasive biomarker in a clinical practice. However, traditional eNoses contain sensors that are notorious for batch-to-batch variation, fouling and aging effects, and sensor drifts.<sup>14,16</sup> The novel measurement of VOCs using physical techniques, coupled with pattern recognition, like FAIMS, have a higher sensitivity and

Received: February 9, 2018

Accepted: June 3, 2018

Published: June 3, 2018

minimal drift. It achieves separation by measuring the differences in the mobility of ionized molecules in high-electric fields.

Data on the potential influence of sampling and storage methods on fecal VOC profiles are scarce. We aimed to evaluate the effects of environmental factors and sampling conditions on fecal VOC profiles using FAIMS. In addition, we aimed to assess under which conditions an optimal diagnostic accuracy could be obtained in the differentiation between pediatric IBD and controls. This may lead to the development of rationale-based standardization protocols on fecal VOC analysis, paving the way toward reliable comparisons between different study outcomes and the implementation of VOC-based diagnostics in clinical practice.

## METHODS

**Study Design.** This case-control study was performed at the outpatient clinic of the pediatric gastroenterology departments in two tertiary referral hospitals, the VU University Medical Center (VUmc) and the Emma Children's Hospital, Academic Medical Center (AMC), both located in Amsterdam, The Netherlands.

**Study Participants. Inflammatory Bowel Disease.** IBD subjects were selected from an existing cohort of de novo treatment-naïve pediatric patients, consisting of 125 subjects (78 CD (Crohn's disease), 47 UC (ulcerative colitis), aged 4 to 17 years, recruited between October 2013 and July 2017 at the VU University Medical Center (VUmc) and Academic Medical Center (AMC). The diagnosis of IBD was based on endoscopic, histologic, and radiologic findings, according to the revised Porto criteria.<sup>17</sup> The localization and behavior of IBD were classified during endoscopy, based on the Paris classification.<sup>18</sup> Physician global assessment (PGA) combined with levels of C-reactive protein (CRP) and fecal calprotectin (FCP) was used as an index of the clinical disease activity.<sup>19,20</sup> All IBD patients were asked to collect a fecal sample prior to endoscopy and bowel preparation.<sup>14</sup> The inclusion criteria also included sufficient fecal material for VOC analysis (3.4 g per subject). Exclusion criteria were the use of antibiotics, probiotics, or immunosuppressive therapy in the three months prior to inclusion, a concomitant diagnosis of a gastrointestinal disease or immunocompromised disease (i.e., HIV, leukemia), and abdominal surgery (except for appendectomy). In addition, children with proven infectious colitis (parasites in stools or positive stool culture for *Salmonella spp.*, *Shigella spp.*, *Yersinia spp.*, *Campylobacter spp.*, or toxigenic *Clostridium spp.*) were excluded.

**Healthy Controls.** Healthy controls (HC) were children 4 to 17 years old selected from elementary and high schools in North-Holland, The Netherlands between June 2016 and December 2016. All participants were asked to collect a fecal sample and complete a questionnaire on abdominal symptoms, bowel habits, including the consistency of their stool using the Bristol stool chart, medication use, and medical history.<sup>21</sup> Exclusion criteria for HC were similar to IBD, with the addition of the diagnosis of IBD and/or a functional gastrointestinal disorder according to the Rome IV criteria based on the questionnaires.

**Matching Procedure.** From the original cohort of 125 IBD patients, 106 were not eligible for this study due to insufficient quantities of the fecal samples. A total of 17 IBD patients (15CD, 2UC) could be matched on age at sample collection and gender with 25 participants in the HC group.

**Ethical Considerations.** This study was approved by the medical ethical review committee (METc) of the VU University Medical Center (VUmc) under file number 2016.393 and by the local medical ethical committee of the Emma Children's Hospital (AMC). Written informed consent was obtained from all parents and from the children when over the age of 12.

**Sample Collection IBD and Controls.** All study participants collected fresh fecal samples in a stool container (Stuhlgefäß 10 mL, Frickenhausen, Germany). Patients with IBD collected their fecal sample prior to endoscopy and bowel lavage. Participants were instructed to store the fecal samples in the freezer at home directly after collection. The samples were transported to the hospital in cooled condition, using cooling elements or ice cubes. Directly upon their arrival in the hospital, the samples were stored in the freezer (−24 °C) until analysis.

**Sample Preparation IBD and Controls.** The influence of fecal sample mass, number of thaw–freeze cycles, and duration of storage at room temperature were assessed by comparing the VOC profiles derived from the subsamples taken from the original fecal sample of each HC and IBD subject. The subsamples were weighed on a calibrated scale (Mettler Toledo, AT 261 Delta Range, Ohio, United States), labeled, and returned to a −24 °C freezer until further handling. We compared the variables of interest with our standard operating procedure (reference standard), defined as a mixture of 500 mg of feces diluted with 10 mL of tap water and kept at room temperature for 10 min prior to analysis. These reference standard settings were chosen since they were used in several previous studies on fecal VOC profiling in a range of gastroenterology diseases and have provided us with positive results.<sup>14,22</sup>

**Variables of Interest.** The effect of fecal sample mass on the diagnostic accuracy of fecal VOC profiles was assessed by comparing subsamples weighing 100 mg and 300 mg with the reference standard mass of 500 mg.

The influence of the number of thaw–freeze cycles on the diagnostic accuracy was analyzed by comparing the reference standard to subsamples, which underwent one, two, and three additional thaw–freeze cycles. For every additional cycle, the sample was kept at room temperature for 10 min and subsequently kept on dry ice until the sample was frozen.

In order to assess the effect of the duration of storage at room temperature on the diagnostic accuracy, VOC from samples kept at room temperature (18 °C) for 180 min were compared to the reference standard. Variables of interest are presented in Table 1.

As described above, the effect of every variable on the diagnostic accuracy of fecal VOCs were assessed by comparing IBD subjects with HC. In addition, we assessed the influence of the variables on the VOC pattern. By combining the HC and IBD subjects, we were able to compare the variables to the reference standard.

**FAIMS Analysis.** For this study a commercially available FAIMS instrument (Lonestar, Owlstone, Cambridge, UK) was used. Prior to the analyses, the FAIMS instrument was checked for contamination using air and water blanks. The fecal samples were thawed to room temperature for 10 min prior to VOC analysis and manually homogenized after diluting the fecal sample with 10 mL of tap water by using a micropipette of 5000  $\mu$ L. The Lonestar was setup as used in previous studies.<sup>14,22,23</sup> To transport the sample headspace into the Lonestar, compressed air (0.1 MPa) was used as the carrier gas. This air meets the European Pharmacopoeia criteria for medical air, and

Table 1. Variables of Interest

variables of interest	fecal sample mass (mg)	thaw–freeze cycles (N)	time out of freezer (min)
reference standard	500	0	10
mass variable 1	300	0	10
mass variable 2	100	0	10
thaw–freeze variable 1	500	1	10
thaw–freeze variable 2	500	2	10
thaw–freeze variable 3	500	3	10
180 min out of freezer	500	0	180
storage time 1	500	0	10
storage time 2	500	0	10

its composition, pressure, temperature, and water density are checked for continuity regularly. When entering the Lonestar, this carrier gas is filtered by a carbon filter (Restek, Bellefonte, PA). The flow rate was set at 2.0 L/min, and the temperature for the sample holder was set at 35 °C, for the lid at 70 °C, and for the filter region at 100 °C. After every sample run, the Lonestar was refreshed using 5 mL of tap water. Furthermore, the dispersion field was set between 0 and 100% (in the ratio of the high electric field to low electric field) and passed through 51 equal settings. The compensation voltage was set between +6 V and –6 V in 512 steps for each dispersion field. All samples were analyzed randomly. Each fecal sample was analyzed three times subsequently, resulting in three matrices and taking 540 s to perform. In order to preclude environmental effects, the first matrix was excluded from analyses since this measurement includes the heaspace gas generated from both the sample and the environment (e.g., air in tubes). For the statistical analysis, only the second matrix was used for optimal diagnostic potential. The third measurement was made as a back-up file, but was not used in this study. The raw data output was analyzed at the School of Engineering, University of Warwick, United Kingdom.<sup>15</sup>

**Statistical Analysis.** The demographic data of each group (IBD patients and healthy controls) were compared using the Mann–Whitney U test for nonparametric continuous data and the Fisher's exact test for dichotomous data using SPSS statistics (version 22, IBM, NY, USA). As previously reported, the FAIMS produces high-dimensional data in terms of the number of features and covariates measured per sample. Therefore, a data compression was performed before feature identification and classification. Each FAIMS data (sample) consists of 52 224 data points in a 2D matrix. Data compression was undertaken by applying a 2D discrete wavelet transformation. For the variables of interest in which the accuracy to discriminate between IBD and HC was assessed, feature selection and classifier training were performed to 90% of data (training set), and class predictions were produced from 10% of the data set (test set) in a 10-fold cross validation. The Wilcoxon rank-sum test was used to calculate *p* values in the training sets to identify which features were best for disease prediction. From this, four statistically important features were used. Four classification algorithms were applied, sparse logistic regression, random forest, Gaussian process, and support vector machine. A receiver operator characteristic (ROC) curve was created to predict the area under the curve (AUC), *p* values, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and

diagnostic accuracy. For the influence of the sampling method on VOC composition, in which IBD and HC samples were combined and measurements of the same subjects samples were repeated, data were analyzed using SPSS statistics 22 (IMB). The raw sensor data was recombined with feature selection using the Wilcoxon rank-sum tests. Paired *t* tests were performed to assess the potential of the features to discriminate between sample handling methods. Scatterplots for the discrimination between samples were created for each variable of interest. Axes depict the recombination of the raw sensor data by means of the features. Individual VOC profiles are illustrated as marked points. The intersection of the lines deriving from the individual VOC profiles demonstrates the mean VOC profile of this specific variable of interest.

**Post Hoc Analyses.** The main target of our study was to assess the optimum sampling method to discriminate between IBD and healthy controls based on VOC analyses by means of FAIMS. We found there is a gap of knowledge on the effects of sample storage duration on VOC integrity. Therefore, the effect of duration of storage in the freezer was analyzed by repeating measurements from a previous study, conducted by van Gaal et al. in which fecal VOC profiles of 36 de novo IBD patients were compared to 24 HC.<sup>14</sup> On the basis of the availability of fecal samples from this study, 10 IBD (all CD) subjects and 10 HC could be included for the reassessment of VOC profiles. Storage time differed by 20 months between the measurements, with a median storage time in the freezer of 43 months for the first and 63 months for the second measurements. For both these analyses, the reference standard was used. Diagnostic accuracy to detect IBD as well as the difference in VOC profile was assessed using the statistical analyses described above.

## RESULTS

**Baseline Characteristics.** Seventeen de novo, treatment-naïve pediatric IBD patients (15 CD, 2 UC) were selected from the original cohort and were matched to 25 HC. Patient characteristics are shown in Table 2. There were no significant differences in age, sex, and sample age between IBD and HC. The range of the sample age was, however, larger in the IBD group compared to HC.

For the assessment of the influence of sample age on diagnostic accuracy, fecal samples of 10 IBD patients (CD only) and 10 HC were selected from the previous study and remeasured.<sup>14</sup> Patients' characteristics for this variable are described in Table 3.

**Fecal VOC Profiles per Variable of Interest.** The results of the VOC analysis displayed per variable of interest are shown in Table 4. For each analysis, the outcome of the sparse logistic regression is noted. A complete overview of the data generated by the four different classification models is given in SI Table 1a–d.

**Standard Operating Procedure.** By the application of the reference standard settings, IBD and HC could be differentiated with high accuracy (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.93, 0.99 (0.96–1), 0.94, 0.96, 0.94, 0.96, and 1.178e-10)(Table 4, SI Table 1a–d, and Figure 1). A typical FAIMS pattern (flame) of both the IBD samples and control samples is depicted in Figure 2.

**Sample Mass.** IBD could be differentiated from HC by using a lower sample mass, but diagnostic accuracy decreased compared to the reference standard for both 300 mg per sample (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.77, 0.69 (0.52–0.86), 0.88, 0.44, 0.52, 0.85, and

Table 2. Baseline Characteristics

	inflammatory bowel disease			<i>p</i> -value
	Crohn's disease ( <i>n</i> = 15)	ulcerative colitis ( <i>n</i> = 2)	healthy controls ( <i>n</i> = 25)	
sex, male ( <i>n</i> , [%])	10 [66.7]	0 [0]	14 [56]	0.858
age, yr (median [IQR])	13.0 [11–15]	[10–16] <sup>c</sup>	12.0 [4.0]	0.614
sample age, mos (median [IQR])	11.0 [2–16]	[11–26] <sup>c</sup>	11.0 [1.0]	0.376
Physician's Global Assessment				
quiescent	1	0		
mild	9	2		
moderate	5	0		
severe	0	0		
fecal calprotectin (μg/g) (median [IQR])	1936 [1006–2390]	[1800–2734] <sup>c</sup>		
CRP (mg/L) (median [IQR])	24.3 [2.5–42]	2.5 <sup>d</sup>		
Crohn's Disease Localization <sup>b</sup>				
ileal (L1)	1			
colonic (L2)	5			
ileocolonic (L3)	6			
proximal disease (L4)	1			
Crohn's Disease Behavior <sup>b</sup>				
B1 (NSNP)	14			
B 1p (NSNP + p)	0			
B2 (S)	1			
B 2p (S + p)	0			
B3 (P)	0			
B 3p (P + p)	1			
Ulcerative Colitis Localization <sup>b</sup>				
proctitis (E1)		1		
left sided (E2)		1		
extensive (E3)		0		

<sup>a</sup>All values were obtained at study inclusion. Localization was obtained by ileocolonoscopy and esophagogastroduodenoscopy before treatment initiation and magnetic resonance enteroclysis. Abbreviations: IQR, interquartile range; NSNP, nonstricturing nonpenetrating; S, stricturing; P, penetrating; p, peri-anal disease. <sup>b</sup>On the basis of the Paris classification for inflammatory bowel disease (24). <sup>c</sup>Minimum-maximum values. <sup>d</sup>One missing value.

0.02101) and 100 mg per sample (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.70, 0.74 (0.59–0.90), 0.76, 0.72, 0.65, 0.82, and 0.00364) (Table 4, SI Table 1a–d, and Figure 1).

**Thaw–Freeze Cycles.** After adding one extra thaw–freeze cycle to the reference standard, a decrease in diagnostic accuracy was observed (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.61, 0.80 (0.65–0.94), 0.76, 0.80, 0.72, 0.83) (Table 4, SI Tables 1–4, and Figure 1). After the addition of a second and third thaw–freeze cycle, differences in the VOC profiles between IBD and HC dissolved (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.64, 0.56 (0.38–0.74), 0.76, 0.48, 0.50, 0.75, and 0.7534 and 0.57, 0.50 (0.32–0.69), 0.47, 0.72, 0.53, 0.67, and 0.5101, respectively) (Table 3, SI Table 1a–d).

**Duration of Storage at Room Temperature.** After keeping the samples at room temperature for 180 min prior to VOC

Table 3. Demographics Sample Analysis of the Influence of Duration Time on VOC Profiles<sup>a</sup>

	Crohn's disease ( <i>N</i> = 10)	healthy controls ( <i>N</i> = 10)	<i>p</i> value
sex, male ( <i>n</i> [%])	5 [50]	2 [20]	0.350
age, yr (median [IQR])	14.1 [3.38]	7.8 [3.72]	0.007
sample age first measurement, mos (median [IQR])	23.4 [21–31]	52.2 [51–52.4] <sup>c</sup>	0.000
sample age second measurement, mos (median [IQR])	43.2 [41–51]	71 [70–72] <sup>c</sup>	0.000
Physician's Global Assessment			
quiescent	0		
mild	0		
moderate	3		
severe	7		
fecal calprotectin (μg/g) (median [IQR])	1067 [1218]		
CRP (mg/L) (median [IQR])	29 [29]		
Crohn's Disease Localization <sup>b</sup>			
ileal (L1)	0		
colonic (L2)	3		
ileocolonic (L3)	7		
proximal disease (L4)	5		
Crohn's Disease Behavior <sup>b</sup>			
B1 (NSNP)	8		
B 1p (NSNP + p)	0		
B2 (S)	0		
B 2p (S + p)	0		
B3 (P)	1		
B 3p (P + p)	1		

<sup>a</sup>All values were obtained at study inclusion. Localization was obtained by ileocolonoscopy and esophagogastroduodenoscopy before treatment initiation and magnetic resonance enteroclysis. Abbreviations: IQR, interquartile range; NSNP, nonstricturing nonpenetrating; S, stricturing; P, penetrating; p, peri-anal disease. <sup>b</sup>On the basis of the Paris classification for inflammatory bowel disease (24). <sup>c</sup>One value missing.

analysis, the differences in VOC outcome between IBD and HC dissolved (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.60, 0.51 (0.32–0.70), 0.59, 0.68, 0.56, 0.71, and 0.4596) (Table 4, SI Table 1a–d).

**Influence of Sampling Method on Overall VOC Composition.** In order to assess the influence of sampling conditions on the detected VOC patterns, HC and IBD subjects were combined to form one single study group. The comparisons between the four features are shown in Table 5. The differences in the VOC pattern between sampling methods are depicted in Figure 3. Both fecal samples weighing 300 mg and 100 mg demonstrated a significantly different VOC profile compared to that of the standard reference mass of 500 mg (features 1, 2, and 4, *p* value <0.001 and for feature 3 *p* value = 0.027 for 300 mg; features 1–4, *p* value <0.0001 for 100 mg). All of the variables in thaw–freeze cycles differed to a similar extent from the reference standard (features 1–4, *p* value <0.0001 for all variables). A similar difference as with the previous variables was seen when comparing the VOC profiles of the reference standard to the VOC profiles of samples kept at room temperature for 180 min prior to VOC analysis (features 1–4, *p* value <0.0001).

**Table 4. Performance Characteristics for the Differentiation between IBD and Healthy for All of the Variables of Interest by Fecal VOC Analysis<sup>a</sup>**

analysis	<i>p</i> value	accuracy	AUC (±95% CI)	cutoff	sensitivity (±95% CI)	specificity (±95% CI)	PPV	NPV
reference standard (17 IBD, 25 HC) <sup>b</sup>	1.178 × 10 <sup>-10</sup>	0.93	0.99 (0.96–1)	0.0014	0.94 (0.71–1)	0.96 (0.8–1)	0.94	0.96
mass variable 1 (17 IBD, 25 HC)	0.02101	0.77	0.69 (0.52–0.86)	0.47	0.88 (0.64–0.99)	0.44 (0.24–0.65)	0.52	0.85
mass variable 2 (17 IBD, 25 HC)	0.003642	0.70	0.74 (0.59–0.9)	0.44	0.76 (0.5–0.93)	0.72 (0.51–0.88)	0.65	0.82
thaw–freeze variable 1 (17IBD, 25HC)	0.0004713	0.61	0.8 (0.65–0.94)	0.49	0.76 (0.5–0.93)	0.8 (0.59–0.93)	0.72	0.83
thaw–freeze variable 2 (17 IBD, 25 HC)	0.7534	0.64	0.56 (0.38–0.74)	0.66	0.76 (0.5–0.93)	0.48 (0.28–0.69)	0.5	0.75
thaw–freeze variable 3 (17IBD, 25HC)	0.5101	0.57	0.5 (0.32–0.69)	0.063	0.47 (0.23–0.72)	0.72 (0.51–0.88)	0.53	0.67
180 min out of freezer (17 IBD, 25 HC)	0.4596	0.60	0.51 (0.32–0.7)	0.13	0.59 (0.33–0.82)	0.68 (0.46–0.85)	0.56	0.71
storage duration, first measurement (10 CD vs 10 HC)	0.0262	0.75	0.75 (0.53–0.97)	0.47	0.7 (0.35–0.93)	0.8 (0.44–0.97)	0.78	0.73
storage duration, second measurement (10 CD vs 10 HC)	0.0376	0.75	0.73 (0.49–0.97)	0.58	0.8 (0.44–0.97)	0.7 (0.35–0.93)	0.73	0.78

<sup>a</sup>For each analysis, the best sparse logistic regression outcome is shown. Sensitivities, specificities, *p* values and AUCs are reported for the respective optimum cut points. Abbreviations: AUC, area under the curve; PPV: positive predictive value; NPV: negative predictive value. <sup>b</sup>Reference standard is defined as a 500 mg sample diluted in 10 mL of water, thawed 10 min to room temperature.

**Post Hoc Analyses.** *Duration of Storage in Freezer.* The diagnostic accuracy to discriminate IBD from controls was not influenced by differences in the duration of storage time prior to VOC analysis (43 versus 63 months) (accuracy, AUC (95% CI), sensitivity, specificity, PPV, NPV, *p* values; 0.75, 0.75 (0.53–0.97), 0.70, 0.80, 0.78, 0.73, 0.0262 versus 0.75, 0.73 (0.49–0.97), 0.80, 0.70, 0.73, 0.78, 0.0376) (Table 4, SI Table 1a–d, and Figure 1). The VOC composition of the two variables showed a significant difference in three features (feature 1–3 with *p* values of <0.0001, < 0.0001, and 0.021, respectively) (Table 5, Figure 3).

## DISCUSSION

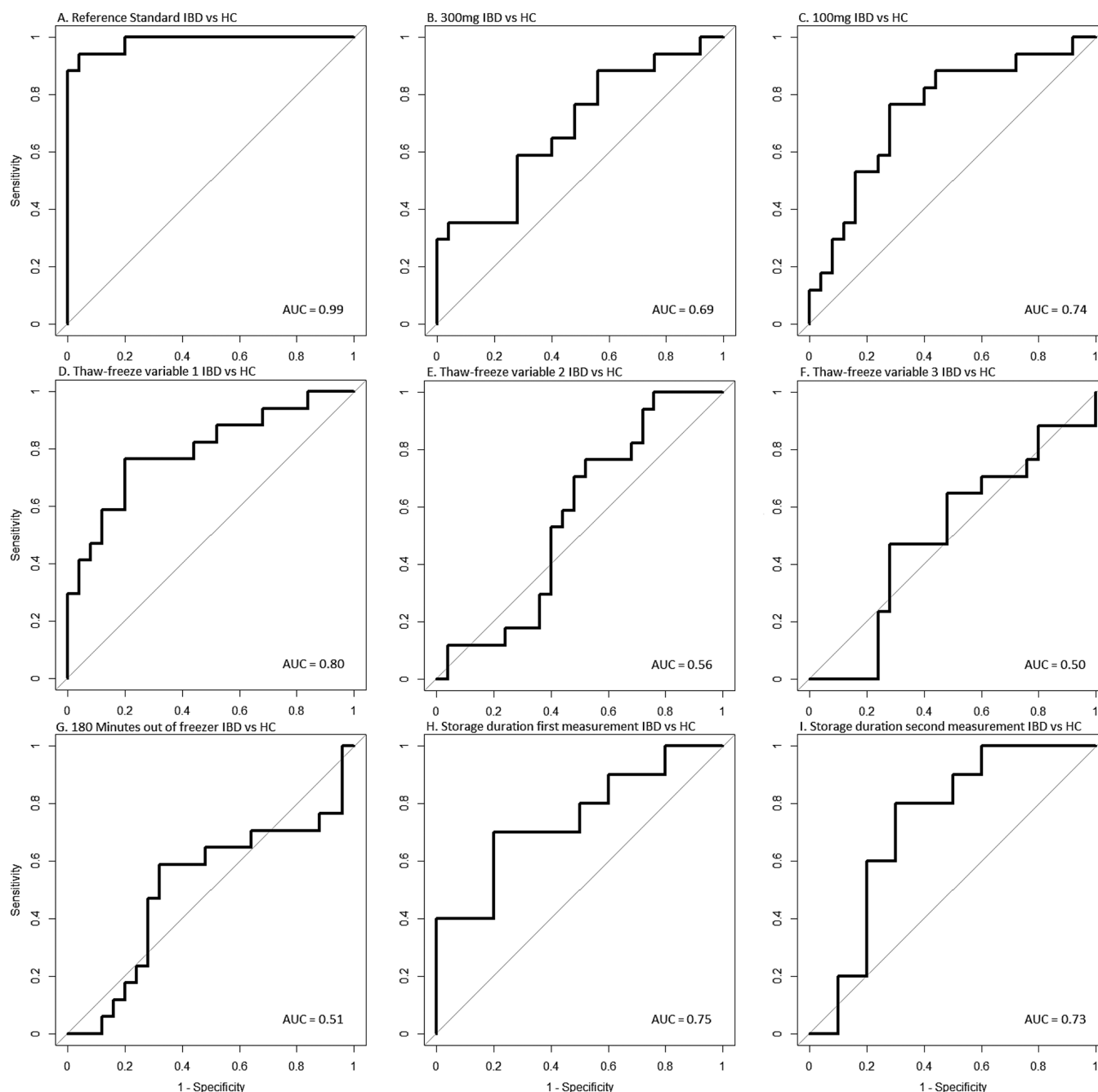
In the present study, VOC profiles and diagnostic accuracy were influenced significantly by altering the sampling conditions. The application of a 500 mg fecal sample mass diluted with 10 mL of tap water and thawed for 10 min prior to analysis after a single thaw–freeze cycle showed the best discriminative accuracy for the differentiation of pediatric IBD and HC.

To our knowledge, this is the first published study to assess under which sampling conditions an optimal accuracy can be obtained in the differentiation between pediatric IBD and healthy controls by analyzing the fecal volatile metabolome using FAIMS. Studies assessing the optimization of sampling methods for fecal metabolome analyses have mainly focused on gas chromatography–mass spectrometry (GC/MS), nuclear magnetic resonance spectroscopy (NMR spectroscopy), and liquid chromatography–mass spectrometry (LC/MS), which are targeted and untargeted methods for the identification of specific metabolites. These studies may hypothetically provide guidance to standardization for the pattern-based FAIMS technique. The results of this study will be discussed and compared to the available literature in the following sections. Regarding sample mass, similar results to eNose, GC/MS, NMR spectroscopy, and LC/MS studies on the fecal metabolome, in both humans and rats, were found in our study, showing a difference between the use of 500 mg from that of lower masses.<sup>16,24,25</sup> Deda and colleagues have shown that the sample weight to volume ratio has a major effect on the number and

signal intensity of features detected in fecal samples with GC/MS. This also applied for the spectral signal intensity when using NMR spectroscopy and for the peak area intensity when using LC/MS.<sup>24</sup> The increased accuracy to differentiate between IBD and HC when using a larger fecal mass, as observed in our study, may be explained by this increase of richness in number and intensity of VOCs.

Observed differences in VOC profiles between fecal samples enduring one versus multiple thaw–freeze cycles are in line with previous research on VOC patterns using different eNose devices.<sup>16,26</sup> It could be hypothesized that these effects are caused by changes in microbiota composition or function; although in a previous study no differences were found in microbiota composition between the analyses of fresh samples versus samples frozen at –80 °C and subsequently thawed prior to analysis.<sup>27</sup> A recent study suggested a release of microbial intracellular contents following thaw–freeze cycles, possibly explaining the effects of thaw–freeze cycles on VOC outcomes.<sup>28</sup> In our study, it was shown that the diagnostic accuracy decreased with the addition of one extra thaw–freeze cycle and that IBD could not be differentiated from HC after the addition of multiple thaw–freeze cycles. Consequently, future studies on fecal VOC should limit the number of thaw–freeze cycles prior to analysis to a maximum of two.

Consistent with the results from a previous study on fecal VOCs using an eNose device, we measured significant differences between fecal VOC profiles measured directly after thawing (as used in the reference standard) and after 180 min stored at room temperature with an accuracy of 0.84.<sup>16</sup> Furthermore, it was demonstrated that the diagnostic accuracy decreased when samples were kept at room temperature for 180 min (AUC = 0.53). These results are in line with a previous study on the impact of storage conditions on crude fecal samples measured by NMR spectroscopy and showed that metabolic variation was influenced by storage at room temperature and at 4 °C.<sup>28</sup> The metabolic profiles of fecal samples did not change after keeping the samples at room temperature for 1 h. However, samples stored for a longer time prior to the analyses gradually shifted. The overall changes that were seen included decreased levels of fumarate, succinate, and glutamate and increased levels

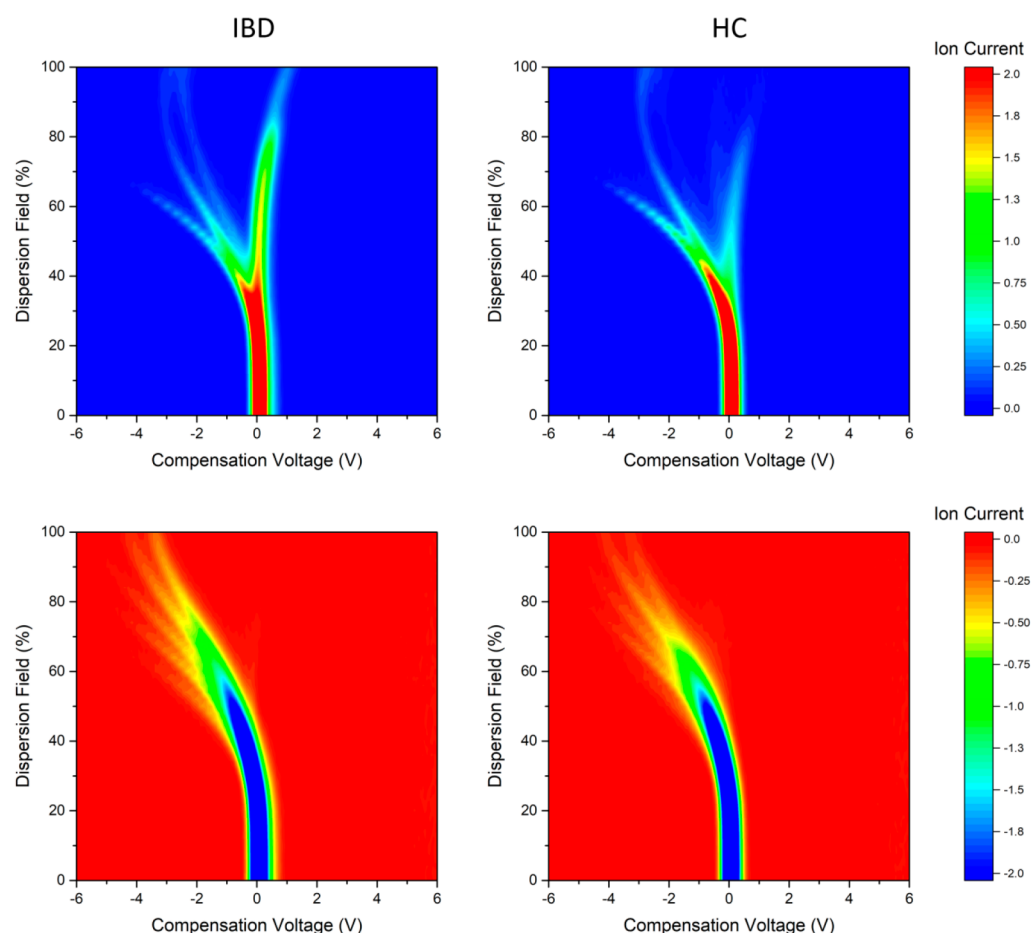


**Figure 1.** Receiver operating characteristics for each variable of interest for the differentiation between inflammatory bowel disease and healthy states. All receiver operating characteristic curves are obtained by sparse logistic regression analyses. Abbreviations: AUC, area under the curve; IBD: Inflammatory bowel disease; HC: Healthy controls.

of methanol, phenylalanine, alanine, and short-chain fatty acids such as acetate, butyrate, propionate, and valerate. To a lesser extent, the same shifts were seen in samples kept at 4 °C, which indicates that the lower temperature slows down the impact on sample integrity, resulting in fewer alterations in the metabolic profile. In another study comparing VOC profiles of fecal samples kept at 1 °C for 14 h prior to GC/MS analysis, there were no significant changes in VOC profiles before and after 14 h.<sup>25</sup> Since the differences between IBD and HC in this study were analyzed by means of pattern recognition, specific metabolic alterations cannot be elucidated in this study. However, it could be hypothesized that the unstable VOC composition when keeping the samples at room temperature is

caused by the ongoing fermentation by the fecal microbiota. Since fermentative processes have shown to be reduced at lower temperature, this could explain why VOC integrity remained stable when samples were kept at 1 and 4 °C in previous studies.<sup>28</sup> Another explanation is the emission of volatiles in the sample and contamination with background volatiles. Fermentation, emission, and contamination could be avoided by measuring the sample directly after collection. However, clinical implementation of VOC analyses would then become a logistics challenge.

Literature on the influence of the storage time of fecal VOCs is scarce. In a study assessing VOC profiles of urine using a similar FAIMS method to the current study, a nine-month shelf life for



**Figure 2.** Typical FAIMS pattern of patients with inflammatory bowel disease and healthy controls. Depicted with a blue background are the positive ion currents. Depicted with a red background are the negative ion currents.

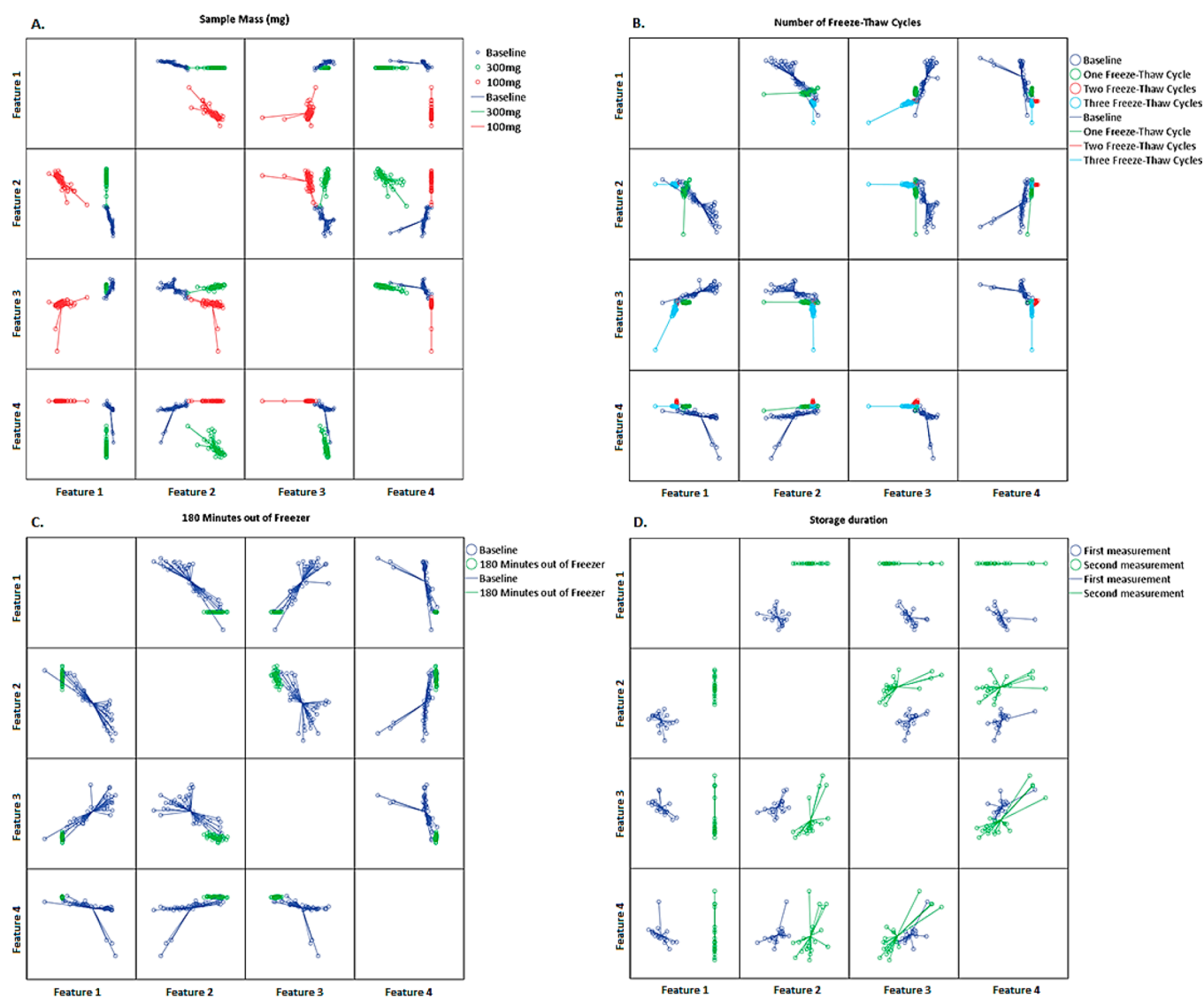
**Table 5. Paired Feature Analyses per Variable of Interest with Corresponding  $p$  Values<sup>a</sup>**

variables of interest	feature 1 ( $p$ value)	feature 2 ( $p$ value)	feature 3 ( $p$ value)	feature 4 ( $p$ value)
	Sample Mass (mg)			
500 vs 300	<0.0001	<0.0001	0.027	<0.0001
500 vs 100	<0.0001	<0.0001	<0.0001	<0.0001
	Number of Freeze–Thaw Cycles			
measured directly vs one cycle	<0.0001	<0.0001	<0.0001	<0.0001
measured directly vs two cycles	<0.0001	<0.0001	<0.0001	<0.0001
measured directly vs three cycles	<0.0001	<0.0001	<0.0001	<0.0001
	Kept at Room Temperature			
180 minutes	<0.0001	<0.0001	<0.0001	<0.0001
	Storage Time			
first vs second measurement	<0.0001	<0.0001	0.021	0.825

<sup>a</sup> $p$  Value <0.05 is considered significant.

urine samples was suggested after it was shown that chemical information was lost over time, regarding both diversity and the concentration of gas emissions.<sup>29</sup> In addition, in a previous study assessing the effect of the sample age on serum, VOCs measured by GC/MS showed that a significant difference in metabolite composition was already seen after being stored for 3 weeks in the freezer.<sup>30</sup> In the current study, fecal VOC profiles for IBD seem less influenced by storage time compared to those in the previous studies on urine and serum, keeping a similar (high) diagnostic accuracy after a storage period of 43 and 63 months. Interestingly, the samples chosen for this comparison were used in a larger study by van Gaal et al. where an area under the curve

of 0.76 was found for a mean storage time of 23 months for the IBD group (25 CD, 21 UC) and 39 months for the HC group.<sup>14</sup> The increase in the AUC of this sub analysis, although analyzed at the same moment, can be explained by the fact that the remaining samples consisted only of CD patients and HC. In the previous study, the AUC for the differentiation between CD and HC was 0.90. There is, however, an important consideration to this post hoc analysis. The diagnostic accuracy was only assessed after a median storage duration of 43 and 63 months. Since there are no previous measurements, it could be possible that massive changes in VOC composition have influenced diagnostic



**Figure 3.** Scatterplot for the differentiation between the sampling methods measured by field asymmetric ion mobility spectrometry, including the (A) sample mass; (B) number of freeze–thaw cycles; (C) 180 min out of freezer; and (D) storage duration. Axes depicted are recombinations of the raw sensor data by means of feature selection using Wilcoxon rank sum analyses, creating four features per measurement. The marked points are the individual VOC signals. The intersection of the lines derived from the individual signals are the mean VOC profile of that specific variable.

accuracy in the initial months after collection. We cannot exclude this influence based on this study.

The main strength of this study is that we used an IBD group and an HC group to assess not only the differences in VOC patterns between sampling methods but also the influence on the diagnostic accuracy for disease detection. In addition, we used the same subjects' samples for each of the analyses, accounting for various confounding factors of influence on fecal VOCs (e.g., smoking habits, medication use, diet). During each experiment, the remaining variables of interest were kept the same, ensuring an optimal comparison based on the variable of interest. Our study also has several limitations. Most importantly, for the influence of storage time on diagnostic accuracy for IBD, we have made use of raw data of a previous study and have reassessed samples with sufficient sample mass. For this analyses we were only able to include CD, and no UC, patients. In addition, sample age differed between groups, which could have influenced diagnostic accuracy by the influence of metabolic degradation on VOC profiles at both measurements.

Second, we have made use of unfiltered and unsterilized tap water for sample dilution and compressed medical air as a carrier gas. This protocol was chosen since it has been found to be a reliable sampling method for the differentiation between various diseases and healthy controls based on fecal VOCs.<sup>31–34</sup> To avoid VOC profile contamination, we have run air and water blanks which were checked on contamination peaks and met the cleanliness criteria. In addition, we have analyzed the samples in a random order and have excluded the first matrix of every sample analysis to avoid air contamination. However, we cannot fully guarantee the exclusion of VOC contamination by differences in tap water composition between measurements. Third, we have not explored the difference between the diagnostic accuracy when using fresh versus frozen samples. As previously described, this seems to be an important influence on urine and serum VOCs. However, since a diagnostic accuracy of 0.99 was found in this study, we believe that freezing our samples has not significantly influenced our study outcomes. Last, it is possible that optimized fecal sampling conditions are



disease-specific, and fecal VOC biomarkers to diagnose IBD might have a different sensitivity to variations in the sampling method compared to fecal VOC biomarkers for other gastrointestinal diseases. We did, however, find significant differences in VOC profiles between sampling methods, emphasizing the importance of the use of one standardized sampling method. Furthermore, it is important to point out that we made use of pattern recognition in this study, which complicates the assessment of the influence of specific metabolites. We have chosen to specifically validate the FAIMS method since this device is an easy-to-use tool which could be suitable for clinical implementation.<sup>35</sup>

This study highlights the need for one standardized methodology, in both a research setting and when using VOC analysis as a (future) clinical tool. On the basis of this and previous study results, we would like to suggest using a standardized protocol with fecal sample masses preferably of 500 mg, no more than one thaw session prior to VOC analysis, and the analysis of samples directly after thawing or, if impossible, keeping the samples frozen until further analysis. Future studies should assess the difference in diagnostic accuracy between fresh samples and frozen samples and the influence of storage duration using multiple-measurement moments after sample collection.

This study showed a high discriminative accuracy to differentiate between IBD and HC when using the standard operating procedure. It was shown that the use of less than 500 mg, multiple thaw–freeze cycles, storage at room temperature, and storage in a freezer all influence the diagnostic accuracy. We therefore suggest to use one standardized protocol when performing fecal VOC analysis. In addition, further studies should focus on finding IBD-specific VOCs to allow for targeted pattern recognition.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.analchem.8b00688](https://doi.org/10.1021/acs.analchem.8b00688).

Performance characteristics with their corresponding area under the curve, sensitivity, specificity, and positive and negative predictive values of fecal volatile organic compound analysis for the differentiation between inflammatory bowel disease and healthy subjects using the supervised classification methods sparse logistic regression, random forest, Gaussian process, and support vector machine. (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Tel: +31 (0)20 444 07 99. Fax: +31 (0)20 444 05 54. E-mail: [s.boschl@vumc.nl](mailto:s.boschl@vumc.nl).

### ORCID

Sofie Bosch: [0000-0001-9202-1674](https://orcid.org/0000-0001-9202-1674)

### Author Contributions

#Shared first, listed alphabetically

### Author Contributions

○Shared last, listed alphabetically

### Author Contributions

T.G.J.M. and N.K.H.B. were the guarantors of this article. S.M.H. and S.B. collected the fecal samples and performed VOC

analysis. J.A.C. and A.N.W. analyzed the results. S.B. drafted the first version of the manuscript. S.M.H., J.A.C., A.N.W., M.K.B., C.J.M., M.A.B., K.H.N.B., and T.G.J.M. reviewed the manuscript for important intellectual content. S.B. finalized the manuscript. All authors agreed to the final version of the manuscript.

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

C.J.J.M. has served as a principal investigator for TEVA Pharma BV. He has received a research grant from TEVA Pharma BV. M.A.B. has served as a speaker for Abbott and FrieslandCampina. He has served as a consultant for Shire, Takeda, Norgine, AstraZeneca, Danone, Friesland Campina, Novalac, and Mead Johnson. N.K.H.B. has served as a speaker for AbbVie and MSD. He has served as a consultant and principal investigator for TEVA Pharma BV and Takeda. He has received a (unrestricted) research grant from Dr. Falk and Takeda. T.G.J.M. has served as a speaker for Mead Johnson. He has received a (unrestricted) research grant from Danone.

## ■ REFERENCES

- (1) Arasaradnam, R. P.; Ouaret, N.; Thomas, M. G.; Qurashi, N.; Heatherington, E.; Nwokolo, C. U.; Bardhan, K. D.; Covington, J. A. *Inflamm Bowel Dis* **2013**, *19*, 999–1003.
- (2) Berkhout, D. J. C.; Niemarkt, H. J.; Benninga, M. A.; ABudding, A. E.; van Kaam, A. H.; Kramer, B. W.; Pantophlet, C. M.; van Weissenbruch, M. M.; de Boer, N. K. H.; de Meij, T. G. J. *Pediatr. Res.* **2018**, *83*, 412.
- (3) Smolinska, A.; Bodelier, A. G.; Dallinga, J. W.; Masclee, A. A.; Jonkers, D. M.; van Schooten, F. J.; Pierik, M. J. *Aliment. Pharmacol. Ther.* **2017**, *45*, 1244.
- (4) Bodelier, A. G.; Smolinska, A.; Baranska, A.; Dallinga, J. W.; Mujagic, Z.; Vanhees, K.; van den Heuvel, T.; Masclee, A. A.; Jonkers, D.; Pierik, M. J.; van Schooten, F. J. *Inflamm Bowel Dis* **2015**, *21*, 1776–1785.
- (5) Arasaradnam, R. P.; McFarlane, M.; Daulton, E.; Skinner, J.; O'Connell, N.; Wurie, S.; Chambers, S.; Nwokolo, C.; Bardhan, K.; Savage, R.; Covington, J. *Dig. Liver Dis.* **2016**, *48*, 148–153.
- (6) Ahmed, I.; Greenwood, R.; Costello, B.; Ratcliffe, N.; Probert, C. S. *Aliment. Pharmacol. Ther.* **2016**, *43*, 596–611.
- (7) Probert, C. S. *Biochem. Soc. Trans.* **2011**, *39*, 1079–1080.
- (8) Nakhleh, M. K.; Amal, H.; Jeries, R.; Broza, Y. Y.; Aboud, M.; Gharra, A.; Ivgi, H.; Khatib, S.; Badarneh, S.; Har-Shai, L.; Glass-Marmor, L.; Lejbkowitz, I.; Miller, A.; Badarny, S.; Winer, R.; Finberg, J.; Cohen-Kaminsky, S.; Perros, F.; Montani, D.; Girerd, B.; Garcia, G.; Simonneau, G.; Nakhoul, F.; Baram, S.; Salim, R.; Hakim, M.; Gruber, M.; Ronen, O.; Marshak, T.; Doweck, I.; Nativ, O.; Bahouth, Z.; Shi, D. Y.; Zhang, W.; Hua, Q. L.; Pan, Y. Y.; Tao, L.; Liu, H.; Karban, A.; Koifman, E.; Rainis, T.; Skapars, R.; Sivins, A.; Ancans, G.; Liepniece-Karele, I.; Kikuste, I.; Lasina, I.; Tolmanis, I.; Johnson, D.; BMillstone, S. Z.; Fulton, J.; Wells, J. W.; Wilf, L. H.; Humbert, M.; Leja, M.; Peled, N.; Haick, H. *ACS Nano* **2017**, *11*, 112–125.
- (9) Broza, Y. Y.; Mochalski, P.; Ruzsanyi, V.; Amann, A.; Haick, H. *Angew. Chem., Int. Ed.* **2015**, *54*, 11036–11048.
- (10) Karban, A.; Nakhleh, M. K.; Cancilla, J. C.; Vishinkin, R.; Rainis, T.; Koifman, E.; Jeries, R.; Ivgi, H.; Torrecilla, J. S.; Haick, H. *Adv. Healthcare Mater.* **2016**, *5*, 2339–2344.
- (11) Bosch, S.; van Gaal, N.; Zuurbier, R. P.; Covington, J. A.; Wicaksono, A. N.; Biezeveld, M. H.; Benninga, M. A.; Mulder, C. J.; de Boer, N. K. H.; de Meij, T. G. J. *Inflamm Bowel Dis* **2018**.
- (12) Probert, C. S.; Reade, S.; Ahmed, I. *Expert Rev. Clin. Immunol.* **2014**, *10*, 1129–1131.
- (13) de Meij, T. G.; de Boer, N. K.; Benninga, M. A.; Lentferink, Y. E.; de Groot, E. F.; van de Velde, M. E.; van Bodegraven, A. A.; van der Schee, M. P. J. *Crohns Colitis* **2014**.

(14) van Gaal, N.; Lakenman, R.; Covington, J.; Savage, R.; de Groot, E.; Bomers, M.; Benninga, M.; Mulder, C.; de Boer, N.; de Meij, T. *J. Breath Res.* **2018**, *12*, 016006.

(15) Covington, J. A.; van der Schee, M. P.; Edge, A. S.; Boyle, B.; Savage, R. S.; Arasaradnam, R. P. *Analyst* **2015**, *140*, 6775–6781.

(16) Berkhout, D.J.; Benninga, M.A.; van Stein, R. M.; Brinkman, P.; Niemarkt, H.J.; de Boer, N. K.; de Meij, T. G. *Sensors* **2016**, *16*, 1967.

(17) Levine, A.; Koletzko, S.; Turner, D.; Escher, J. C.; Cucchiara, S.; de Ridder, L.; Kolho, K. L.; Veres, G.; Russell, R. K.; Paerregaard, A.; Buderus, S.; Greer, M. L.; Dias, J. A.; Veereman-Wauters, G.; Lionetti, P.; Sladek, M.; Martin de Carpi, J.; Staiano, A.; Ruemmele, F. M.; Wilson, H. D. C. *J. Pediatr. Gastroenterol. Nutr.* **2014**, *58*, 795–806.

(18) Levine, A.; Griffiths, A.; Markowitz, J.; Wilson, D. C.; Turner, D.; Russell, R. K.; Fell, J.; Ruemmele, F. M.; Walters, T.; Sherlock, M.; Dubinsky, M.; Hyams, J. S. *Inflamm Bowel Dis* **2011**, *17*, 1314–1321.

(19) Turner, D.; Otley, A. R.; Mack, D.; Hyams, J.; de Bruijne, J.; Uusoue, K.; Walters, T. D.; Zachos, M.; Mamula, P.; Beaton, D. E.; Steinhart, A. H.; Griffiths, A. M. *Gastroenterology* **2007**, *133*, 423–432.

(20) Hyams, J. S.; Ferry, G. D.; Mandel, F. S.; Gryboski, J. D.; Kibort, P. M.; Kirschner, B. S.; Griffiths, A. M.; Katz, A. J.; Grand, R. J.; Boyle, J. T.; et al. *J. Pediatr. Gastroenterol. Nutr.* **1991**, *12*, 439–447.

(21) Lewis, S. J.; Heaton, K. W. *Scand. J. Gastroenterol.* **1997**, *32*, 920–924.

(22) Bomers, M. K.; Menke, F. P.; Savage, R. S.; Vandenbroucke-Grauls, C. M.; van Agtmael, M. A.; Covington, J. A.; Smulders, Y. M. *Am. J. Gastroenterol.* **2015**, *110*, 588–594.

(23) Covington, J. A.; Wedlake, L.; Andreyev, J.; Ouaret, N.; Thomas, M. G.; Nwokolo, C. U.; Bardhan, K. D.; Arasaradnam, R. P. *Sensors* **2012**, *12*, 13002–13018.

(24) Deda, O.; Chatziioannou, A. C.; Fasoula, S.; Palachanis, D.; Raikos, N.; Theodoridis, G. A.; Gika, H. G. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2017**, *1047*, 115–123.

(25) Reade, S.; Aggio, A.; Khalid, R.; Pritchard, T.; Ewer, D. M.; Probert CS, A. K. *J. Anal. Bioanal. Tech.* **2014**, *5*, 1000184.

(26) Chan, D. K.; Leggett, C. L.; Wang, K. K. *World J. Gastroenterol* **2016**, *22*, 1639–1649.

(27) Fouhy, F.; Deane, J.; Rea, M. C.; O'Sullivan, O.; Ross, R. P.; O'Callaghan, G.; Plant, B. J.; Stanton, C. *PLoS One* **2015**, *10*, e0119355.

(28) Gratton, J.; Phetcharaburanin, J.; Mullish, B. H.; Williams, H. R.; Thursz, M.; Nicholson, J. K.; Holmes, E.; Marchesi, J. R.; Li, J. V. *Anal. Chem.* **2016**, *88*, 4661–4668.

(29) Esfahani, S.; Sagar, N.M.; Kyrou, I.; Mozdiak, E.; O'Connell, N.; Nwokolo, C.; Bardhan, K.D.; Arasaradnam, R.P.; Covington, J.A. *Biosensors* **2016**, *6*, 4.

(30) Forbes, S. L.; Rust, L.; Trebilcock, K.; Perrault, K. A.; McGrath, L. T. *Forensic Sci., Med., Pathol.* **2014**, *10*, 570–582.

(31) Sofie, B.; Zuurbier, Roy P.; Covington, James A.; Wicaksono, Alfian N.; Biezeveld, Maarten H.; Benninga, Marc A.; Mulder, Chris J.; de Boer, Nanne K.H.; de Meij, Tim G.J. *Inflammatory Bowel Diseases* **2018**, *1*.

(32) van Gaal, N.; Lakenman, R.; Covington, J.; Savage, R.; de Groot, E.; Bomers, M.; Benninga, M.; Mulder, C.; de Boer, N.; de Meij, T. *J. Breath Res.* **2018**, *12*, 016006.

(33) Berkhout, D. J. C.; Niemarkt, H. J.; Benninga, M. A.; Budding, A. E.; van Kaam, A. H.; Kramer, B. W.; Pantophlet, C. M.; van Weissenbruch, M. M.; de Boer, N. K. H.; de Meij, T. G. *J. Pediatr. Res.* **2018**, *83*, 412–419.

(34) Berkhout, D. J. C.; Niemarkt, H. J.; Buijck, M.; van Weissenbruch, M. M.; Brinkman, P.; Benninga, M. A.; van Kaam, A. H.; Kramer, B. W.; Andriessen, P.; de Boer, N. K. H.; de Meij, T. G. *J. Pediatr. Gastroenterol. Nutr.* **2017**, *65*, e47–e52.

(35) Arasaradnam, R.; Wicaksono, A.; O'Brien, H.; Kocher, H.M.; Covington, J.A.; Crnogorac-Jurcevic, T. *Gastroenterology* **2018**, *154*, 485.