Exhaled volatile organic compounds discriminate patients with chronic obstructive pulmonary disease from healthy subjects

Vasiliki Besa, Helmut Teschler, Isabella Kurth, Amir Maqbul Khan, Paul Zarogoulidis, Joerg Ingo Baumbach, Urte Sommerwerck, Lutz Freitag, Kaid Darwiche

Abstract: Chronic obstructive pulmonary disease (COPD) is a chronic airway inflammatory disease characterized by incompletely reversible airway obstruction. This clinically heterogeneous group of patients is characterized by different phenotypes. Spirometry and clinical parameters, such as severity of dyspnea and exacerbation frequency, are used to diagnose and assess the severity of COPD. The purpose of this study was to investigate whether volatile organic compounds (VOCs) could be detected in the exhaled breath of patients with COPD and whether these VOCs could distinguish COPD patients from healthy subjects. Moreover, we aimed to investigate whether VOCs could be used as biomarkers for classifying patients into different subgroups of the disease. Ion mobility spectrometry was used to detect VOCs in the exhaled breath of COPD patients. One hundred and thirty-seven peaks were found to have a statistically significant difference between the COPD group and the combined healthy smokers and nonsmoker group. Six of these VOCs were found to correctly discriminate COPD patients from healthy controls with an accuracy of 70%. Only 15 peaks were found to be statistically different between healthy smokers and healthy nonsmokers. Furthermore, by determining the cutoff levels for each VOC peak, it was possible to classify the COPD patients into breathprint subgroups. Forced expiratory volume in 1 second, body mass index, and C-reactive protein seem to play a role in the discrepancies observed in the different breathprint subgroups.

Keywords: breath analysis, COPD, ion mobility spectrometry, volatile organic compounds

Introduction

Chronic obstructive pulmonary disease (COPD) is a common preventable and treatable disease worldwide and is one of the leading causes of morbidity and mortality. COPD is characterized by persistent limitation of airflow due to an abnormal inflammatory response of the lungs to the inhalation of noxious particles and gases and is mainly caused by cigarette smoking. The limitation of airflow is best measured by spirometry, which is widely available and reproducible. However, the lung function parameters obtained from spirometry do not always yield assessments of airway inflammation and severity of disease. Dyspnea and rate of exacerbations are essential for investigating the underlying mechanisms of the disease and for assessing severity of the disease and the response to the therapy. Furthermore, some subjects are not compliant enough to provide a good spirometry test; therefore, there is a need for additional methods that provide more information about the disease than airway limitation.

Noninvasive methods, including the analysis of exhaled breath, have been studied in the past decades for their applicability in the assessment of airway inflammation and as possible diagnostic tools in several inflammatory lung diseases. A large number of biomarkers in breath have been investigated as possible indicators of inflammation,
to diagnose and monitor the disease as well as to evaluate the response to treatment. Exhaled breath is known to contain many volatile organic compounds (VOCs) at very low concentrations (nanomolar [10^{-9}] or picomolar [10^{-12}]). These VOCs may be products of various inflammatory and metabolic processes, either physiological or disease-related, that take place in the airways and other parts of the human body, or products of the oxidative stress that occurs in diseases such as asthma and COPD. Therefore, the detection of VOCs in the exhaled breath is an attractive method to investigate possible biomarkers for diagnosing, monitoring, and assessing the oxidative stress of various lung diseases. Previous studies have shown that a profile of VOCs could be used as a biomarker in lung cancer, asthma, sarcoidosis, tuberculosis, and obstructive sleep apnea syndrome. Profiles of exhaled VOCs provide fingerprints of diseases, the so-called “breathprints,” allowing for discriminating between patients and healthy subjects, without the need for chemical identification of the underlying substances.

Nowadays, many techniques are available to collect and analyze VOCs, such as the electronic nose (eNose), gas chromatography (GC) followed by either mass spectrometry (GCMS) or flame ionization detection (GC-FID), proton transfer reaction mass spectrometry (PTR-MS), selected ion flow tube mass spectrometry (SIFT-MS), laser spectroscopy, colorimetric sensor array, and gold nanoparticle sensors (GNPs). Exhaled VOCs detected by gas chromatography–mass spectrometry have been shown to discriminate patients with asthma or COPD from healthy subjects; as well as patients with COPD from patients with asthma or lung cancer with the use of the electronic nose. Ion mobility spectrometry (IMS) has been used in several studies for detecting VOCs in patients with several lung diseases. IMS is a new, on-site method, rapid and easy to use for detecting and separating VOCs according to their detection time, drift time, and their concentration in the exhaled breath. Moreover, IMS enables the visualization of VOCs in a three-dimensional topography, the so-called IMS chromatograph.

The purpose of this study was to detect in the exhaled breath of patients with COPD VOCs that could be used as biomarkers in the diagnosis of the disease. Furthermore, we aimed to detect VOCs that could be used as possible biomarkers for classifying patients into different subtype groups of the disease.

**Methods**

**Study subjects**

The patients included in the study were recruited from the Department of Pneumology, Ruhrlandklinik, University Hospital of Essen, Germany. Patients had an established diagnosis of COPD according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines. All patients had a history of smoking (>20 pack per years) and an irreversible limitation of airflow (reversibility <12% predicted forced expiratory volume in 1 second [FEV1] and <200 mL in FEV1) after inhalation of β2-agonist. Patients were asked for comorbidities and in case of other existing respiratory and inflammatory diseases or malignancies, they were excluded from the study. All patients had no signs of acute exacerbation for at least 4 weeks prior to enrollment.

The control groups consisted of healthy smokers and nonsmokers, and all were employees of the hospital. All healthy subjects had no history of respiratory disease and underwent lung function measurements to exclude obstructive lung disease. The nonsmoking subjects had a smoking history of less than two packs per year and had stopped smoking for at least 1 year. The study was approved by the ethics committee of the University of Essen and all subjects provided written informed consent.

**Study design**

All VOC measurements were performed with the same ion mobility spectrometer device, in the same room and at the same daytime to avoid biased variations. The study participants were requested to refrain from eating, drinking, and smoking for 2 hours prior to the measurement. All subjects underwent a medical history review, a physical examination, and a collection of breath air. Repeated measurements on the same day were performed in a subgroup of the healthy subjects to assess the within-day repeatability of the measurements.

**Lung function tests**

The lung function tests were performed by a trained lung function technician using Zan500 Body (nSpire Health, Oberthulba, Germany) according to the American Thoracic Society/European Respiratory Society recommendations.

**Collection of exhaled breath**

All subjects were requested to exhale through a mouthpiece connected via a Teflon tube with the spectrometer. In each case, end-tidal breath, controlled by a flow sensor, was collected in a sample loop of 10 mL in volume. The sample air was collected and transferred to a multicapillary column for a first chromatographic separation. Using the software VOCan 1.3 (B&S Analytik, Dortmund, Germany), the dead volume was adjusted and fixed to 500 mL. The expiration was controlled by an ultrasound CO2 sensor element.
Analysis of exhaled breath

The IMS coupled to the multicapillary column (MCC/IMS) was a BioScout (B&S Analytik), consisting of the MCC/IMS and a SpiroScout (GanhornMedizin Electronic, Niederlauer, Germany) as a sample inlet unit. The major parameters of this setup are summarized elsewhere.24,26 In this spectrometer, a 550 MBq $^{60}$Ni $\beta$-radiation source was applied for the ionization of the carrier gas (air). The spectrometer was connected to a polar, multicapillary column (MCC, type OV-5; Multichrom Ltd, Novosibirsk, Russia) that was used as the preseparation unit. In this MCC, the analytes of exhaled breath were sent through 1,000 parallel capillaries, each with an inner diameter of 40 μm and a film thickness of 200 nm. The total diameter of the separation column was 3 mm. The relevant MCC parameters are listed in Table 1.

Data mining and evaluation

The peaks were characterized using the software Visual Now (B&S Analytik), which is described elsewhere.27,28 All peaks found are characterized by their positions with respect to drift time (corresponding $1/K_0$ value), retention time, and their concentration in relation to the peak height.22,28

Statistical analysis

The statistical analysis was performed using SPSS (version 17.0). The normality of distributions was evaluated with the Kolmogorov–Smirnov test. The comparisons among groups were performed with one-way analysis of variance for normally distributed data and with the Kruskal–Wallis test for skewed data, with appropriate post hoc tests to adjust for multiple comparisons (Bonferroni). The difference of a numerical variable between two groups was evaluated with an unpaired $t$-test or Mann–Whitney $U$-tests for normal and skewed data, respectively.

Furthermore, a classification model was constructed using twofold cross-validation. The training data were randomly split in two parts: one to develop the model and one to measure its performance.

Results

The subjects’ characteristics are summarized in Table 2. Breath analysis was performed on 45 COPD patients (aged 56.2±8.5 years), 23 healthy smokers (aged 38.7±14 years) and 28 healthy nonsmokers (aged 42.5±8.4 years). There were no significant differences in sex and body mass index (BMI) among the three groups. The COPD patients had a heavier smoking history than the healthy smokers (33±14 packs per year vs 18.6±18 packs per year).

The mean FEV$_1$ (predicted percentage) in patients with COPD was 28.3±18.9 vs 105.6±8.0 in healthy smokers and 97±14 in healthy nonsmokers. The severity of COPD was GOLD IV (n=21), GOLD III (n=16), GOLD II (n=5), and GOLD I (n=3). Patients were referred to the Department of Pneumology, University Hospital of Essen, to examine the potential for endoscopic or surgical lung volume reduction or for their addition to the lung transplant waiting list.

Table 1 Characteristics of ion mobility spectrometer

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Spectrometer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ionization source</td>
<td>$^{60}$Ni (555 MBq) $\beta$-radiation</td>
</tr>
<tr>
<td>Electric field strength</td>
<td>320 V/cm</td>
</tr>
<tr>
<td>Length of drift region</td>
<td>12 cm</td>
</tr>
<tr>
<td>Diameter of drift region</td>
<td>15 mm</td>
</tr>
<tr>
<td>Length of ionization chamber</td>
<td>15 mm</td>
</tr>
<tr>
<td>Shutter opening time</td>
<td>300 μs</td>
</tr>
<tr>
<td>Shutter impulse time</td>
<td>100 ms</td>
</tr>
<tr>
<td>Drift gas</td>
<td>Synthetic air (20.5% O$_2$ 79.5% N$_2$)</td>
</tr>
<tr>
<td>Drift gas flow</td>
<td>100 mL/min</td>
</tr>
<tr>
<td>Temperature</td>
<td>Ambient temperature</td>
</tr>
<tr>
<td>Pressure</td>
<td>101 kPa (ambient pressure)</td>
</tr>
<tr>
<td>MCC type</td>
<td>OV-5, polar</td>
</tr>
<tr>
<td>Column temperature</td>
<td>40°C</td>
</tr>
</tbody>
</table>

Abbreviation: MCC, multicapillary column.

Table 2 Patients’ characteristics

<table>
<thead>
<tr>
<th>Parameters</th>
<th>COPD (n=45)</th>
<th>Healthy smokers (n=23)</th>
<th>Healthy nonsmokers (n=28)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>56.2±8.5$^{a,b}$</td>
<td>38.7±14</td>
<td>42.5±8.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sex (female/male)</td>
<td>27/18</td>
<td>14/9</td>
<td>14/14</td>
<td>ns</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>25.7±6.8</td>
<td>25.4±4.6</td>
<td>25.4±4</td>
<td>ns</td>
</tr>
<tr>
<td>Ex-smokers/current smokers</td>
<td>40/5$^{a,b}$</td>
<td>12/11</td>
<td>0</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Smoking history (pack per years)</td>
<td>33±14$^a$</td>
<td>18.6±18</td>
<td>0</td>
<td>ns</td>
</tr>
<tr>
<td>FEV$_1$ (% pred)</td>
<td>28.3±18.9$^{a,b}$</td>
<td>105.6±8</td>
<td>97±14</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FEV$_1$ (L)</td>
<td>0.89±0.73$^{a,b}$</td>
<td>3.4±0.72</td>
<td>3.4±0.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FEV$_1$/VC (pred)</td>
<td>35.3±20.5$^{a,b}$</td>
<td>80.4±5</td>
<td>75.3±4</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Notes: *Statistically significant difference compared to healthy smokers; $^{a,b}$statistically significant difference compared to healthy nonsmokers.

Abbreviations: FEV$_1$, forced expiratory volume in 1 second; VC, vital capacity; BMI, body mass index; COPD, chronic obstructive pulmonary disease; ns, not significant; pred, predicted.
A total of 224 VOC peaks were identified, as characterized by drift and retention times (Figure 1). The signal intensity of the peaks was statistically evaluated with the Wilcoxon rank sum test. One hundred and thirty-seven peaks were found to have a statistically significant difference between the COPD group and the combined healthy smokers and nonsmokers group. When the healthy smokers and the healthy nonsmokers were compared with each other, only 15 peaks were found to be statistically different between the two groups (EES, EJ, FI, FL, FP, GF, P, P5 22, PS3, P1, P2, P27, P38, P4, and P40) (Supplementary information). Using 50:50 cross-validation, six peaks (PS0, PS14, PS47, P51, P52, and P2) from among the total of 137 (Figure 2) were found to correctly classify (with an accuracy of 71%, 70%, 70%, 71%, 70%, and 67%, respectively) the COPD patients from the healthy controls (smokers and nonsmokers combined).

For each peak, a cutoff value in signal intensity was determined such that the COPD patients were correctly classified as COPD and the healthy controls as non-COPD, with a specificity of 100%. Using this cutoff value for each peak, we examined if the COPD patients lie above or below this cutoff value. It becomes apparent that the vast majority of the patients (41 out of 45) have at least five out of the six peaks above or below this cutoff. When their peaks were found above the cutoff, it was defined as “typical COPD breathprint.” When the peaks were found to be below (again at least five of the six peaks), patients were defined as having “no typical COPD breathprint.”

In only four patients, no more than four peaks were in accordance to classify the patients into one of these two breathprint subgroups. These four patients were excluded from further analysis, which was performed to investigate the reason for these discrepancies. These four patients seemed to have no common characteristics in their breathprints. Thus, 23 of 41 patients were classified as having a “typical COPD breathprint” and 18 as not having a “typical COPD breathprint.”

To further investigate the possible reasons for the discrepancies in the breathprints of these patients, we examined a set of clinical parameters that were likely to discriminate these two breathprint subgroups. From these parameters, the subgroup of patients with a “typical COPD breathprint” by VOC profile had statistically significant higher FEV₁, BMI, and lower CRP in comparison with the subgroup without “typical COPD breathprint” by VOC profile (Table 3). For other parameters, no significant differences between the two subgroups could be found.

Discussion

Using IMS, we have shown that the VOC profile of the exhaled breath of COPD patients is different from that of healthy subjects. Six VOCs were found to classify and distinguish correctly the COPD patients from healthy subjects. Further, by using cutoff levels for each of the six VOC peaks, it was possible to discriminate the COPD patients into two breathprint subgroups, patients with a “typical COPD breathprint” and patients without a “typical COPD breathprint.”

The use of noninvasive techniques for assessing airway inflammation, such as the analysis of exhaled breath, has developed rapidly since nitric oxide was described as an important biomarker in the exhaled breath of asthma patients. Apart from nitric oxide, a large number of biomarkers have been tested in exhaled breath with various noninvasive methods, such as induced sputum and exhaled breath condensate, to investigate, assess, and monitor airway inflammation or oxidative stress in lung diseases. These noninvasive diagnostic methods are very attractive for clinical settings also because they are easy to perform and safe for the patients. Furthermore, these noninvasive methods, in contrast to invasive methods, can be easily repeated and applied by patients with severe diseases.

Recent studies have also shown discrimination of COPD patients from healthy subjects using several methods for detecting VOCs in exhaled breath. Our findings are in accordance with the recent study of Basanta et al which discriminated COPD patients from healthy subjects with an accuracy of 69% using gas chromatography–time-of-flight mass spectrometry. In a study by van Berkel et al using gas chromatography–mass spectrometry, six peaks were detected that classified correctly 92% of the patients with a sensitivity of 98% and a specificity of 88%. However, these two studies included a larger proportion of current smokers (31% and 76%, respectively) than was included in our study (13%). This higher proportion of smokers in these two studies could possibly explain the higher accuracy of discriminating COPD patients from healthy subjects compared to the results of our study. Phillips et al also showed using gas chromatography–mass spectrometry, in a large study group of 119 COPD patients and 63 healthy subjects, that VOCs can distinguish COPD patients from healthy subjects with a high accuracy of 74%. In their study, the COPD and control group were matched for age and BMI because these factors could possibly affect VOCs. In our study, no difference was found for BMI between the two groups, but age was statistically different between the two groups. While age is a possible factor affecting VOCs via oxidative stress, currently only few data provide support for this suggestion.

The study of Phillips et al implies that smoking status influences the correct classification of COPD patients as...
Exhaled VOCs discriminate patients with COPD from healthy subjects.

**Figure 1** The IMS chromatogram showing the 224 peaks that were detected in the exhaled breath of COPD patients.

**Abbreviations:** IMS, ion mobility spectrometry; COPD, chronic obstructive pulmonary disease.
Active smoking may be a confounding factor in the analysis of exhaled breath. Previous studies have shown that active smoking influences the VOC profile in exhaled breath.\(^8,36,37\) However, when we compared healthy smokers with healthy nonsmokers in our study, 15 peaks that were found to distinguish the two groups were not similar to the peaks that distinguished COPD patients from healthy subjects (smoking and nonsmoking combined). Consequently, it is probable that smoking did not affect our results.

Identification of the underlying substances detected in the exhaled breath was beyond the scope of this investigation. Substance identification is not strictly necessary because IMS allows for the detection of VOC profiles that are themselves useful as potential markers for detecting disease without the need for further chemical analysis. A recent study could show using the electronic nose that VOCs could identify bacterial colonization in COPD patients.\(^38\) This could be a subject for a future study using the IMS. An interesting aspect of this study was that by identifying cutoff values for each of the six peaks that discriminated COPD patients from healthy subjects, these peaks were found to be unequally distributed among COPD patients. In contrast, COPD patients could be further classified into two breathprint subgroups: those with a “typical COPD breathprint” and those without a “typical COPD breathprint.” This finding implies that there are plausible subtypes within COPD patients and that these subtypes can be identified by breath analysis. This result is in accordance with the study of Basanta et al\(^31\) which

![Graph showing box-and-whisker plots for six peaks](image)

**Figure 2** Box-and-whisker plots of the six peaks P2, P51, P52, Ps0, Ps14, and Ps47 that differentiated the COPD patients from the healthy subjects (healthy smokers and healthy nonsmokers).

**Abbreviation:** COPD, chronic obstructive pulmonary disease.

### Table 3 Clinical parameters of COPD patients with and without “typical COPD breathprint”

<table>
<thead>
<tr>
<th>Clinical parameters</th>
<th>Patients with “typical COPD breathprint” (n=23)</th>
<th>Patients without “typical COPD breathprint” (n=18)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEV(_1) (% pred)</td>
<td>36.8±21</td>
<td>18.7±9</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>27.6±8</td>
<td>22.8±4</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>CRP (mg/dL)</td>
<td>0.6±1</td>
<td>1.7±2</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Age (years)</td>
<td>58±11</td>
<td>54±6</td>
<td>0.140</td>
</tr>
<tr>
<td>Serum creatinine (mg/dL)</td>
<td>0.82±0.2</td>
<td>0.78±0.2</td>
<td>0.789</td>
</tr>
<tr>
<td>Pack per years</td>
<td>32.6±18</td>
<td>32±11</td>
<td>0.946</td>
</tr>
<tr>
<td>pO(_2) (mmHg)</td>
<td>70.5±10</td>
<td>75.5±15</td>
<td>0.486</td>
</tr>
<tr>
<td>pCO(_2) (mmHg)</td>
<td>42.7±7</td>
<td>47.7±9</td>
<td>0.115</td>
</tr>
</tbody>
</table>

**Abbreviations:** FEV\(_1\), forced expiratory volume in 1 second; BMI, body mass index; CRP, C-reactive protein; COPD, chronic obstructive pulmonary disease; pred, predicted.
showed that VOCs in exhaled breath identified clinically relevant subgroups, eg, COPD patients with sputum eosinophilia, asymptomatic smokers, or patients with frequent exacerbations.

It is known that COPD is a complex and multidimensional disease with different pulmonary and extrapulmonary manifestations.59 Although diagnosis and monitoring of the disease are based on lung function tests, it is now known that the measurement of FEV1 alone does not express the full complexity of the disease. Various phenotypes reflect several clinical, immunological, and inflammatory mechanisms of COPD that influence the outcome of the disease and the response to treatment.33,40

To elucidate the parameters that are associated with the discrepancy in classifying COPD patients by the VOC profiles identified in our study, we analyzed further the differences in FEV1, BMI, CRP, age, serum creatinine, packs per year, pO2, and pCO2 between the two subgroups. Our finding that patients with lower FEV1 do not show a “typical COPD breathprint” may be explained by higher oxidative stress in a more severe stage of the disease. This suggestion could explain the finding that patients without “typical COPD breathprint” had lower BMI and higher CRP, indicators for systemic inflammation, which is more prominent in severe COPD. Another possible explanation could be that metabolic processes independent of inflammation influence the breathprint of patients with more severe disease in a way that they have similarities with the breathprint of healthy subjects. Further characterization of these COPD patients using other inflammatory biomarkers or clinical data is of interest and should be investigated in further studies.

Our study has certain limitations. Repeated IMS measurements were performed in healthy subjects on the same day, showing good reproducibility. However, reproducibility was not tested in COPD patients. It is suggested that sample variability and short-term effects of practice or exertion should be considered in breath analysis tests.41 Incalzi et al3 suggest that VOC patterns are reproducible in healthy subjects and patients with very severe COPD, whereas these are less reproducible in COPD patients with less severe disease. This finding may reflect hypoxemia, which characterizes these patients. As the majority of the patients measured in this study suffer from severe COPD, variability of IMS measurement might not be a confounding factor in our study. Another limitation of our study is that no information was collected regarding medication of the patients. Further studies are needed to test the possible effects of medication on exhaled breath and to test repeatability and reproducibility in COPD patients. Although identification of the underlying substances of the detected VOCs was beyond the interest of this study, this could be a subject of future studies that would probably reveal the origin of these compounds and further elucidate the pathophysiological mechanisms involved in the pathogenesis of COPD.

**Conclusion**

In this study, it was shown that the identification of VOCs using IMS was able to distinguish COPD patients from healthy subjects. Two subgroups of COPD patients were identified according to their VOC breathprints. Further studies with larger sample size are needed to completely characterize these subgroups, as well as to identify the underlying substances of the VOCs.

**Acknowledgments**

The authors thank Gabriele Kentrat (Ruhrlandklinik, Essen, Germany) for her assistance with data collection. This research has been supported by the “Deutsche Forschungsgemeinschaft” within the Collaborative Research Center (“Sonderforschungsbereich”) SFB 876.

**Disclosure**

Baumbach JI has shares as inventor in diverse patents and is shareholder of a company that produces ion mobility spectrometers. The other authors report no conflicts of interest in this work.

**References**


