Exhaled breath analysis for the diagnosis of pneumonia

van Oort, P.M.P.

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses): Other

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
EXHALED BREATH METABOLOMICS REVEALS A PATHOGEN-SPECIFIC RESPONSE IN A RAT MODEL FOR STREPTOCOCCUS PNEUMONIAE AND PSEUDOMONAS AERUGINOSA PNEUMONIA


American Journal of Physiology – Lung Cellular and Molecular Physiology, 2019
DOI: 10.1152/ajplung.00449.2018
ABSTRACT

Volatile organic compounds (VOCs) in breath can reflect host and pathogen metabolism and might be used to diagnose pneumonia. We hypothesized that rats with Streptococcus pneumoniae (SP) or Pseudomonas aeruginosa (PA) pneumonia can be discriminated from uninfected controls by thermal desorption – gas chromatography – mass-spectrometry (TD-GC-MS) and selected ion flow tube – mass spectrometry (SIFT-MS) of exhaled breath.

Male adult rats (n=50) received an intra-tracheal inoculation of 1) 200 µL saline, 2) 1x10^7 colony forming units (CFU) of SP or 3) 1x10^7 CFU of PA. 24 hours later the rats were anaesthetized, tracheotomized and mechanically ventilated. Exhaled breath was analyzed via TD-GC-MS and SIFT-MS. Area under the receiver operating characteristic curves (AUROCCs) and correct classification rate (CCRs) were calculated after leave-one-out cross-validation of sparse partial least squares-discriminant analysis (sPLS-DA).

Analysis of GC-MS data showed an AUROCC (95% CI) of 0.85 (0.73 – 0.96) and CCR of 94.6% for infected vs. non-infected animals, AUROCC 0.98 (0.94 – 1) and CCR of 99.9% for SP vs. PA, 0.92 (0.83 – 1.00) and CCR of 98.1% for SP vs. controls and 0.97 (0.92 – 1.00) and CCR of 99.9% for PA vs. controls. For these comparisons the SIFT-MS data showed AUROCCs of 0.54, 0.89, 0.63 and 0.79, respectively.

Exhaled breath analysis discriminated between respiratory infection and no infection, but with even better accuracy between specific pathogens. Future clinical studies should not only focus on the presence of respiratory infection, but also on the discrimination between specific pathogens.
Exhaled breath metabolomics for pneumonia in a rat model

INTRODUCTION

Exhaled breath analysis of volatile organic compounds (VOCs) represents a promising new technique for diagnosing respiratory infection\textsuperscript{1–3}. Our recent review\textsuperscript{4}, however, has shown that current studies using breath analysis did not show sufficient diagnostic accuracy and lack consistency to be used for pneumonia in mechanically ventilated intensive care unit (ICU) patients.

Studies investigating individual infection related VOCs or VOC patterns in human breath encounter certain challenges, such as: 1) all possible pathogens are investigated at once; 2) for pneumonia no gold standard is available\textsuperscript{5}; and 3) due to co-existing factors such as comorbidities, drugs, and diet, it might be difficult to determine the biochemical origin of VOCs. The application of exhaled breath metabolomics or breathomics is rapidly expanding\textsuperscript{6,7}. Specific VOC profiles for certain bacterial strains can be identified\textsuperscript{8}. In vitro studies using bacterial cultures\textsuperscript{9,10} do not take into account the host response, and bacteria appear to grow differently in culture media compared to living lung tissue\textsuperscript{11}.

To date animal studies investigating VOCs for diagnosis of pneumonia\textsuperscript{12–14} primarily used secondary electrospray ionization - mass spectrometry (SESI-MS) as analytical platform for breath analysis, resulting in breathprint patterns associated with certain microorganisms. However, identification of specific individual VOCs is preferable, since this could guide future human studies. Capture of breath on suitable sorbent tubes followed by thermal desorption into gas chromatography-mass spectrometry (TD-GC-MS) can identify individual VOCs and is currently seen as the gold standard regarding exhaled breath analysis\textsuperscript{11}. Selected ion flow tube – mass spectrometry (SIFT-MS) offers the possibility of on-line breath analysis, and thus might enable future application for exhaled breath monitoring at the patient’s bedside.

Within the scope of this study exhaled breath in a rat pneumonia model was investigated, for two common causative pathogens of pneumonia: Streptococcus pneumoniae (SP) and Pseudomonas aeruginosa (PA). It was hypothesized that 1) rats with SP or PA pneumonia can be discriminated from uninfected controls; and 2) the different pathogens can be distinguished using exhaled breath analysis.
Chapter 6

METHODS

The study was approved by the Animal Welfare Body at the AMC Amsterdam, the Netherlands (project number LEICA125AD-1).

Experimental groups

Male adult specific pathogen-free Sprague-Dawley rats \((n=50)\) weighing \(~350\) grams (Envigo, Netherlands) received an intra-tracheal inoculation of either: 1) a total of \(1 \times 10^7\) colony forming units (CFU) of SP (ATCC 6303; Rockville, USA) \((n=18)\); or 2) a total of \(1 \times 10^7\)CFU of PA (PA103; Iglewski Laboratory, USA) \((n=16)\), under light anaesthesia using isoflurane 3%; or 3) 200µL saline \((n=16)\) for the control group.

Anaesthesia and mechanical ventilation

24 hours post-inoculation, an anaesthetic mixture (0.15mL/100g body weight) of 1.8mL ketamine (100mg/mL; Eurovet Animal Health, Netherlands), 0.5mL dexmedetomidine (0.5mg/mL; Vetoquinol, Netherlands), 0.2mL atropine (0.5mg/mL; Eurovet Animal Health) and 0.5mL NaCl 0.9% was injected. The rats were weighed, tracheotomised and connected to a mechanical ventilator (Dräger, Netherlands). The rats were pressure controlled ventilated with 16cm\(\text{H}_2\text{O}\) over 2cm\(\text{H}_2\text{O}\) positive end-expiratory pressure, using a fraction of inspired oxygen of 32%.

Exhaled breath collection

For breath sampling, a stainless steel tube filled with sorbent material (for GC-MS: Tenax\textsuperscript{TM} GR 60/80; Interscience, Netherlands; and for SIFT-MS Carbograph 1TD/Carbopack X; Markes International, UK) was inserted between the expiratory ventilator tubing and a pump (Markes). For 10 minutes VOCs were absorbed onto the steel sorbent tube with a flow of 100mL/min. The sorbent tubes were stored at 4°C for a maximum of 14 days until analysis.

Other samples

Directly after collection of the exhaled breath samples, the rats were sacrificed. For the bronchoalveolar lavage (BAL) sample, three 2mL aliquots of saline were instilled and directly withdrawn from the right lung. The upper lobe of the left lung was fixed in 4% buffered formaldehyde for later paraffin embedding, sectioning and staining at the pathology department. The middle and lower lobes of the left lung were homogenized.
**Thermal desorption gas chromatography–mass spectrometry**

Sorbent tubes were placed within a TD unit (TD100; Markes) and heated (250°C for 15min, flow 30mL/min). The VOCs were captured on a cold trap (5°C), which was rapidly heated to 300°C for 1min, after which the molecules were splitless injected through a transfer line at 120°C onto an Inertcap 5MS/Sil GC column (30m, diameter 0.25mm, film thickness 1μm, 1,4-bis(dimethylsiloxy)phenylene dimethyl polysiloxane; Restek, Netherlands) at 1.2mL/min. The oven temperature was isothermal at 40°C for 5min, then increased to 270° at 10°C/min and kept isothermal at 270°C for 5min.

Molecules were ionized using electron ionization (70eV), and the fragment ions were detected using a quadrupole mass spectrometer (GCMS-GP2010; Shimadzu, Netherlands) with a scan range of 37–300Da. Ion fragment peaks were used for statistical analysis. The predictive fragment ions were manually checked in the raw chromatograms and corresponding metabolites were tentatively identified using National Institute of Standards and Technology library (NIST, Gaithersburg, USA); we followed the Metabolomics Standards Initiative for metabolite identification\(^\text{15}\).

**Thermal desorption selected ion flow tube–mass spectrometry**

The discriminatory power of the GC-MS and SIFT-MS full-scan VOC patterns was compared. SIFT-MS (Voice200; Syft Technologies) was used as an off-line instrument in combination with a TD unit (UNITY; Markes). A full scan was performed in the mass-to-charge (m/z) ratio of 15+ to 200+, without the limitation of changing VOC levels throughout breathing manoeuvres, as would be the case when analysing on-line. Sorbent tubes were placed in an autosampler (ULTRA; Markes) connected to the TD unit. TD was performed in tube conditioning mode and the tubes were heated to 270°C (flow 30mL/min) for 10min. The VOCs were recollected in a 1L Tedlar® gas sampling bag (Sigma-Aldrich) at the split outlet. The Tedlar® bag was placed at the sample inlet of the SIFT-MS (Voice200; Syft Technologies, New Zealand) and full scan was initiated with a scan range from m/z 15+ to 200+ for 3 precursor ions (H\(_2\)O+, NO+, O\(_2\)+), a dwell time of 100ms, a count limit of 10000 and 8 repeats. Raw data in counts/second of all scanned ions were corrected for the instrument calibration function (ICF) of the measurement day. The ion counts were multiplied by the ion-specific instrument calibration function. The ICF-corrected data were then used for statistical analysis.
Chapter 6

Infection assessment

Serial 10-fold dilutions of the homogenized lung and the BAL fluid were plated on blood agar plates and incubated overnight at 37°C. The number of CFUs were counted the next morning. Cell counts in the BALF were measured (Z2 Coulter Particle Counter; Beckman Coulter Corporation, USA) and neutrophils counted (Cytospin™ 4 Cytocentrifuge; Thermo Scientific™, USA).

Histologic examination of the 4µm hematoxylin and eosin-stained lung sections was performed by a pathologist blind to group identity. Lung inflammation and damage was determined using a lung infection scoring system as described previously.

Data analysis

All statistical analyses were performed in R statistics through the R-studio interface. A p-value ≤0.05 was considered statistically significant for single comparisons. P-values were corrected for multiple-testing by Benjamini-Hochberg correction. Diagnostic accuracy was measured by the area under the receiving operating characteristics curve (AUROCC).

The allocation of an animal to pneumonia or control group was the primary dependent variable. All analyses were repeated for SP vs. control, PA vs. control and SP vs. PA, to study the inter-pathogen variance. The VOCs measured by TD GC-MS and SIFT-MS were used as 2 separate predictor matrices for pneumonia status.

First, high dimensional datasets with VOCs were reduced by principal component (PC) analysis. The first 6 PCs were retained, capturing 57% of variance. A conservative number of PCs was used because of the relatively low number of animals. Mann-Whitney U test was used to test differences in PCs between groups. PCs with a p-value ≤0.1 were used for logistic regression (LR) analysis. Second, individual VOCs were compared using the “limma” package and p-values and fold changes were reported and shown in a volcano plot. VOCs with an adjusted p-value ≤0.05 were identified. Third, sparse partial least square discriminant analysis (sPLS-DA; MixOmics package) with leave-one-out cross-validation was used to identify the most discriminatory VOCs and estimate the accuracy of such a selected dataset. We could not use bootstrap analyses due to low sample number so we employed leave-one-out where data from an individual animal was left
out of the modelling. The correct classification rate (CCR) was calculated by comparing the AUROCC of the leave-one-out cross-validated model to a similarly constructed model for 1000 randomly permuted labels, as is recommended\textsuperscript{20}.

**RESULTS**

All animals survived the 24h post-inoculation and the 1-hour period of mechanical ventilation. Median BALF white cell count was (in cells/mL) $13.8 \times 10^5$ (IQR: $8.7 \times 10^5$–$16.7 \times 10^5$) for the SP rats, $5.9 \times 10^5$ (IQR: $4.0 \times 10^5$–$11.2 \times 10^5$) for the PA rats and $1.3 \times 10^5$ (IQR: $1.2 \times 10^5$–$1.5 \times 10^5$) for the control rats ($p<0.001$). The CFU counts of the BALF samples differed significantly between the groups: no CFUs were seen on the agar plates for BALF of the PA and control rats, compared to a median of $4.8 \times 10^6$ (IQR: $1.2$–$8.8 \times 10^6$) CFU/mL for the SP animals ($p<0.001$). Only the homogenate of the SP group showed significant growth ($p<0.001$: $1.0 \times 10^9$ (IQR: $7.4$–$1.0 \times 10^9$) CFU/mL), compared to 650 (IQR: $0$–$4.4 \times 10^3$) CFU/mL for the PA rats and 0 (IQR: $0$–$1.4 \times 10^3$) CFU/mL for the controls. Microscopic counts of the percentages of neutrophils present on the stained cytospin preparations differed between groups ($p<0.001$), with a median of 88.5 (IQR: 72.5–95.3) for the SP animals, 81 (IQR: 68.5–89) for the PA group and 2.5 (IQR: 0–5) for the controls.

The percentage of pneumonia on histopathological investigation was significantly higher in the SP rats ($p<0.001$). Pneumonia scores were significantly higher in the infected vs. the non-infected animals: median pneumonia score was 8 (IQR: 6–10.5) for the SP rats and 5.5 (IQR: 3–6.5) for the PA rats, compared to 3 (IQR: 2–4) for the controls ($p<0.001$).

**Thermal desorption gas chromatography-mass spectrometry**

The analysis of significant PCs (using PCs 1, 4 and 5) and subsequent LR model for infected vs. non-infected animals showed an AUROCC of 0.93 (95%-CI: 0.85–1). The AUROCC (using PC 1, 4 and 5) was 0.93 (95%-CI: 0.84–1) for SP vs. controls, 0.98 (95%-CI: 0.94–1) for PA vs. controls using PC 4 and 5, and 0.99 (95%-CI: 0.97–1) for SP vs. PA using PC 1, 3 and 5.

Figure 1 shows the group comparisons. Comparing infected vs. non-infected animals, 16% of VOCs were significantly different between groups, resulting in a false discovery rate (FDR) of 31.3%. For SP vs. controls the significant rate was 30% (FDR 16%), for PA
vs. controls 15% (FDR 33%) and for SP vs. PA 42% (FDR 12%). Table 1 shows identified VOCs, with an adjusted $p$-value of <0.05 to limit chances of false discovery.

**Figure 1.** Volcano plots for the group comparisons.

sPLS-DA with leave-one-out cross-validation at the animal level followed by LR showed an AUROCC of 0.85 (95%-CI: 0.73–0.96) for infected vs. non-infected animals, with a correct classification rate (CCR) of 94.6% (Figure 2a). SP vs. controls had an AUROCC of 0.92 (95%-CI: 0.83–1) (CCR 98.1%), PA vs. controls an AUROCC of 0.97 (95%-CI: 0.92–1) (CCR 99.9%), and SP vs. PA an AUROCC of 0.98 (95%-CI: 0.94–1) (CCR 99.9%) (Figure 3a).

**Figure 2.**

SPLS-DA analysis with leave-one-out cross-validation: infected (purple triangles: SP; purple dots: PA) vs. non-infected (green rhombus) animals: 2a. (left); GC-MS results; 2b. (right): SIFT-MS results.
Figure 3. SPLSDA analysis with leave-one-out cross-validation: SP (red triangles) vs. PA (blue dots) animals: 3a. (left); GC-MS results; 3b. (right): SIFT-MS results.

Table 1.

<table>
<thead>
<tr>
<th>Group comparison</th>
<th>VOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infected vs. non-infected</td>
<td>Octane, 4-methyl-</td>
</tr>
<tr>
<td></td>
<td>Octane, 2-5-dimethyl</td>
</tr>
<tr>
<td></td>
<td>Unidentified naphthalene compound</td>
</tr>
<tr>
<td></td>
<td>Unidentified cyclic compound</td>
</tr>
<tr>
<td></td>
<td>Unidentified</td>
</tr>
<tr>
<td></td>
<td>Unidentified branched aldehyde</td>
</tr>
<tr>
<td></td>
<td>Tetra chloroethylene</td>
</tr>
<tr>
<td></td>
<td>Unidentified cyclic compound</td>
</tr>
<tr>
<td>SP vs. control</td>
<td>Octane, 4-methyl-</td>
</tr>
<tr>
<td></td>
<td>Octane, 2-5-dimethyl</td>
</tr>
<tr>
<td></td>
<td>Unidentified naphthalene compound</td>
</tr>
<tr>
<td></td>
<td>Unidentified</td>
</tr>
<tr>
<td></td>
<td>Unidentified cyclic compound</td>
</tr>
<tr>
<td></td>
<td>Hexadecane</td>
</tr>
<tr>
<td></td>
<td>Unidentified</td>
</tr>
</tbody>
</table>
Table 1 (continued).

<table>
<thead>
<tr>
<th>Group comparison</th>
<th>VOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP vs. control (continued)</td>
<td>Unidentified</td>
</tr>
<tr>
<td></td>
<td>Hexane, 2-,4-dimethyl-</td>
</tr>
<tr>
<td></td>
<td>2-Propanol, 1-methyloxy-</td>
</tr>
<tr>
<td></td>
<td>Nonane, 2-methyl-</td>
</tr>
<tr>
<td></td>
<td>Heptane, 2-,4-dimethyl</td>
</tr>
<tr>
<td></td>
<td><strong>Unidentified cyclic compound</strong></td>
</tr>
<tr>
<td>PA vs. control</td>
<td><strong>Unidentified branched aldehyde</strong></td>
</tr>
<tr>
<td></td>
<td>2-propenoic acid, 2-ethylhexyl ester</td>
</tr>
<tr>
<td></td>
<td>Unidentified cyclic compound</td>
</tr>
</tbody>
</table>

Identified VOCs with an adjusted $p$-value of <0.05, per comparison. In bold: identical VOCs showing overlap between the group comparisons.

**Thermal desorption selected ion flow tube-mass spectrometry**

The analyses were repeated for the SIFT-MS data. For infected vs. non-infected animals the significant PCs (PC 1 and 4) had an AUROCC of 0.78 (95%-CI: 0.62–0.94). For SP vs. controls the AUROCC (using PC 1, 2 and 4) was 0.82 (95%-CI: 0.67–0.96), for PA vs. controls the AUROCC was 0.85 (95%-CI: 0.69–1) using PC 4, and for the SP vs. PA animals the AUROCC was 1.0 (95%-CI: 1–1) using PC 1 and 2.

Aforementioned method for sPLSDA analysis resulted in an AUROCC of 0.54 (95%-CI: 0.38–0.71) for infected vs. non-infected animals (Figure 2b) (CCR 1.6%), an AUROCC of 0.63 (95%-CI: 0.43–0.83) (CCR 26.9%) for SP vs. controls, an AUROCC of 0.79 (95%-CI: 0.62–0.96) (CCR 77.6%) for PA vs. controls, and an AUROCC of 0.89 (95%-CI: 0.77–1) (CCR 19.6%) for SP vs. PA (Figure 3b).

**DISCUSSION**

The exhaled breath of rats with SP or PA pneumonia can be discriminated from uninfected controls with good accuracy using GC-MS. The discriminative accuracy was even higher for the discrimination between the two specific pathogens. Overall, GC-MS results provided better results than SIFT-MS as analytical platform for this purpose.
This is the first study that demonstrates an evidently better discriminative performance of breath analysis when used for discrimination between pathogens instead of distinguishing healthy from diseased. So far, clinical studies have been aiming to show a potential for breath analysis to diagnose a variety of lung diseases, e.g. ARDS\textsuperscript{21} and COPD\textsuperscript{22}. Clinical studies investigating breathomics for the diagnosis of respiratory infection, showed a general focus on the identification of distinctive individual VOCs or breathprints to be served as biomarkers for pneumonia\textsuperscript{1–3}, and not specifically for the causative pathogens. In contrast, our results demonstrate that breath analysis can differentiate bacteria with a higher diagnostic accuracy. In retrospect, this finding seems to be more in line with the available in-vitro data. A meta-analysis of all available studies linked more VOCs to one or a few pathogens and rarely found VOCs in the headspace of all studies\textsuperscript{11}.

Among the identified VOCs were several alkane hydrocarbons (Table 1). Alkanes are associated with oxidative stress\textsuperscript{10}, yet have been linked to pneumonia as well\textsuperscript{4}. The abundance of octane may be secondary to peroxidation of oleic acid\textsuperscript{10}. The other identified hydrocarbons – hexadecane (previously linked to lung infection\textsuperscript{12}), 2-,4-dimethylhexane, 2-methylnonane and 2-,4-dimethylheptane (previously associated with S. aureus and E. coli infection\textsuperscript{23}) – were mainly produced by SP. 2-Propanol is – as endogenous compound – suggested to be a product of an enzyme mediated reduction of acetone\textsuperscript{24} and, like octane, might serve as a possible biomarker\textsuperscript{25}. Tetrachloroethylene is used primarily in the dry cleaning industry and likely to be a contaminant. 2-Propenoic acid is known to derive from ventilator and tubing\textsuperscript{25}. Table 1 shows that presently many of our discovered VOCs could not be named and remained unidentified, which does not limit them to be of value, for their specific combination of retention time and mass spectrum enables future recognition of these markers in clinical studies and therefore they might still serve as markers for the presence of a specific bacterium.

Animal models provide a controlled environment free of genetic or behavioural influences, allowing selected pathogens to be studied without coexisting microorganisms or diseases contaminating the breath signal. Several studies in murine models focused on the differentiation between individual pathogens by detecting selective VOC patterns\textsuperscript{13,26}. Since the present study used GC-MS, individual VOCs could be identified as opposed to the recognition of patterns. These VOCs could serve as specific markers for particular pathogens and could thus be applied for future human exhaled breath studies\textsuperscript{27}.
the diagnostic accuracy of single markers provides less accuracy than composite signals. Pathogen identification by VOC analysis in exhaled breath may be most feasible by breathprint analysis and not solely by the analysis of one or several specific VOCs.

A strength of this study is the controlled environment of the established animal model, using a breath sampling technique that had been proven successful in rat experiments. Genetically identical rats were used and a precisely regulated amount of bacteria was inoculated. Another strength is the use of two independent analytical platforms that showed similar trends in results. Limitations of the study were the small panel of pathogenic bacteria that was studied and the relatively limited amount of animals that was used. Due to the small number of animals used in these experiments, cross-validation had to be performed at the leave-one-animal-out level. Another limitation is the number of VOCs of interest that remained unidentified.

To date, GC-MS is seen as the gold standard for exhaled breath analysis. SIFT-MS has the advantage of being quick (few minutes), without requiring calibration standards for the measured VOCs. Furthermore it can be used as an on-line instrument enabling real-time measurements, without the need of sample preconcentration. However, an off-line approach was used in the current study, involving a rather novel variation of coupling a TD unit upfront the instrument, as earlier described for detection of selected compounds in ambient air. In this off-line confirmation, full scan mode is more feasible: a chosen range of ions with defined m/z ratio can be scanned for a chosen time, without the limitation of on-line sampling, including changing VOC levels throughout breathing manoeuvres. An additional advantage of using SIFT-MS off-line in combination with the TD unit is the possibility to preconcentrate and potentially measure trace elements in exhaled breath which would fall below the detection limit without preconcentration.

In the present study, both the GC-MS and the SIFT-MS technique delivered adequate accuracies regarding the ability of VOCs to differentiate between causative pathogens, but only GC-MS could discriminate between infected and non-infected rats. GC-MS data for infected vs. non-infected animals could have been over-fit, as indicated by the high FDRs in the univariate analysis. Nevertheless, GC-MS results have proved superior to SIFT-MS results before in gaseous samples containing large numbers of VOCs at high concentrations.
In conclusion, the current focus of exhaled breath metabolomics might have to be reconsidered: in addition to the aim to detect the general presence of respiratory infection, clinical studies should concentrate more on the discrimination between pathogens.
REFERENCES


Exhaled breath metabolomics for pneumonia in a rat model


