Exhaled breath metabolomics reveals a pathogen-specific response in a rat pneumonia model for two human pathogenic bacteria: a proof-of-concept study.

Pouline M. van Oort[1]*[‡](MD MPhil), Paul Brinkman[1] (PhD), Gitte Slingers[2] (MSc), Gudrun Koppen[3] (PhD), Adrie Maas[1], Joris J. Roelofs[1] (MD PhD), Ronny Schnabel[4] (MD PhD), Dennis C. Bergmans[4] (MD PhD), M. Raes[2] (MD PHD), Roy Goodacre[5] (MD PhD), Stephen J. Fowler[6] (MD PhD), Marcus J. Schultz[1] (MD Prof) and Lieuwe D. Bos[1] (MD PhD), on behalf of the BreathDx Consortium[+]

*Correspondence: p.m.vanoort@amc.uva.nl

[‡]Location of the performed laboratory experiments

[1]Department of Intensive Care, Amsterdam UMC – location Academic Medical Centre (AMC), Amsterdam, the Netherlands

[2]Hasselt University, Hasselt, Belgium

[3]Flemish Institute for Technological Research, Mol, Belgium

[4]Maastricht University Medical Centre (MUMC), Maastricht, the Netherlands

[5]Manchester Institute of Biotechnology, Manchester, United Kingdom

[6]University of Manchester, Manchester, United Kingdom

[+]A full list of the members of the BreathDx Consortium: Waqar M. Ahmed, Antonio Artigas, Jonathan Bannard-Smith, Lieuwe D. J. Bos, Marta Camprubi, Luis Coelho, Paul Dark, Alan Davie, Emili Diaz, Gemma Goma, Timothy Felton, Stephen J. Fowler, Royston Goodacre, Hugo Knobel, Oluwasola Lawal, Jan-Hendrik Leopold, Tamara M.E. Nijsen, Pouline M. P. van Oort, Pedro Povoa, Craig Johnson, Nicholas J. W.

Rattray, Guus Rijnders, Marcus J. Schultz, Ruud Steenwelle, Peter J. Sterk, Jordi Valles, Fred Verhoeckx, Anton Vink, Hans Weda, Iain R. White, Tineke Winters, Tetyana Zakharkina.

Financial disclosure: no (external) funding to be reported.

Key words: exhaled breath analysis; pneumonia; infection; biomarkers

To be submitted as a Rapid Report

Word count (including figure legends and references): 3995

Abstract

Introduction: 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.

Methods: 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).

Results: 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.

Discussion: 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.

Introduction

Exhaled breath analysis of volatile organic compounds (VOCs) represents a promising new technique for diagnosing respiratory infection (12, 20, 24). Our recent review(19), 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(13); 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(6, 18). Specific VOC profiles for certain bacterial strains can be identified(5). *In vitro* studies using bacterial cultures(16, 17) do not take into account the host response, and bacteria appear to grow differently in culture media compared to living lung tissue(8).

To date animal studies investigating VOCs for diagnosis of pneumonia(1, 28, 29) 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(8). 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.

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 ~$1x10^7$ colony forming units (CFU) of *SP* (ATCC 6303; Rockville, USA) (*n*=18)*;* or 2) a total of ~$1x10^7$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 16cmH_2O over 2cmH_2O 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^{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(26).

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_3O+, 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.

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(4).

Data analysis

All statistical analyses were performed in *R statistics* through the R-studio interface(22). A p-value ≤0.05 was considered statistically significant for single comparisons. *P*-values were corrected for multiple-testing by Benjamini-Hochberg correction(2). 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(14). 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 permutated labels, as is recommended(27).

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) $13.8x10^5$ (IQR: $8.7x10^5$–$16.7x10^5$) for the *SP* rats, $5.9x10^5$ (IQR: $4.0x10^5$–$11.2x10^5$) for the *PA* rats and $1.3x10^5$ (IQR: $1.2x10^5$–$1.5x10^5$) 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 $4.8x10^6$ (IQR: 1.2–8.8 $x10^6$) CFU/mL for the *SP* animals ($p<0.001$). Only the homogenate of the *SP* group showed significant growth ($p<0.001$; $1.0x10^9$ (IQR: 7.4–1.0 $x10^9$) CFU/mL), compared to 650 (IQR: 0–$4.4x10^3$) CFU/mL for the *PA* rats and 0 (IQR: 0–1.4^10^3) 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

188 (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
189 controls.

190 The percentage of pneumonia on histopathological investigation was significantly higher in the
191 *SP* rats (p<0.001). Pneumonia scores were significantly higher in the infected *vs*. the non-infected
192 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,
193 compared to 3 (IQR: 2–4) for the controls (p<0.001).

194

195 *TD GC-MS*

196 The analysis of significant PCs (using PCs 1, 4 and 5) and subsequent LR model for infected *vs*. non-
197 infected animals showed an AUROCC of 0.93 (95%-CI: 0.85–1). The AUROCC (using PC 1, 4 and 5) was
198 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
199 0.99 (95%-CI: 0.97–1) for *SP* vs. *PA* using PC 1, 3 and 5.

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

205 sPLSDA with leave-one-out cross-validation at the animal level followed by LR showed an
206 AUROCC of 0.85 (95%-CI: 0.73–0.96) for infected vs. non-infected animals, with a correct classification
207 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*
208 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:
209 0.94–1) (CCR 99.9%)(Figure 3a).

TD SIFT-MS

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(10) and COPD(3). 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*(12, 20, 24), and not specifically for *the*

232 *causative pathogens*. In contrast, our results demonstrate that breath analysis can differentiate bacteria
233 with a higher diagnostic accuracy. In retrospect, this finding seems to be more in line with the available
234 *in-vitro* data. A meta-analysis of all available studies linked more VOCs to one or a few pathogens and
235 rarely found VOCs in the headspace of all studies(8).

236 Among the identified VOCs were several alkane hydrocarbons (Table 1). Alkanes are associated
237 with oxidative stress(16), yet have been linked to pneumonia as well(19). The abundance of octane may
238 be secondary to peroxidation of oleic acid(16). The other identified hydrocarbons – hexadecane
239 (previously linked to lung infection(28)), 2-,4-dimethylhexane, 2-methylnonane and 2-,4-
240 dimethylheptane (previously associated with *S. aureus* and *E. coli* infection(11)) – were mainly produced
241 by *SP*. 2-Propanol is – as endogenous compound – suggested to be a product of an enzyme mediated
242 reduction of acetone(25) and, like octane, might serve as a possible biomarker(7). Tetrachloroethylene is
243 used primarily in the dry cleaning industry and likely to be a contaminant. 2-Propenoic acid is known to
244 derive from ventilator and tubing(7). Table 1 shows that presently many of our discovered VOCs could
245 not be named and remained *unidentified*, which does not limit them to be of value, for their specific
246 combination of retention time and mass spectrum enables future recognition of these markers in clinical
247 studies and therefore they might still serve as markers for the presence of a specific bacterium.

248 Animal models provide a controlled environment free of genetic or behavioural influences,
249 allowing selected pathogens to be studied without coexisting microorganisms or diseases contaminating
250 the breath signal. Several studies in murine models focused on the differentiation between individual
251 pathogens by detecting selective VOC patterns(29, 30). Since the present study used GC-MS, individual
252 VOCs could be identified as opposed to the recognition of patterns. These VOCs could serve as specific
253 markers for particular pathogens and could thus be applied for future human exhaled breath studies(21).
254 However, 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(9). 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(8). 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(23). 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(15).

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.

Conflict of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

290 Figure legends

291

292

293 Figure 1. Volcano plots for the group comparisons.

294

295

296 Figure 2. SPLSDA analysis with leave-one-out cross-validation: infected (purple triangles: *SP*; purple dots:

297 *PA*) vs. non-infected (green rhombus) animals: 2a. (left); GC-MS results; 2b. (right): SIFT-MS results.

298

299 Figure 3. SPLSDA analysis with leave-one-out cross-validation: *SP* (red triangles) vs. *PA* (blue dots)

300 animals: 3a. (left); GC-MS results; 3b. (right): SIFT-MS results.

301

302

303 References

304 1. **Bean HD**, **Jimenez-Diaz J**, **Zhu J**, **Hill JE**. Breathprints of model murine bacterial lung infections are
305 linked with immune response. *Eur Respir J* 45: 181–190, 2015.

306 2. **Benjamini Y**, **Hochberg Y**. Controlling the false discovery rate: a practical and powerful approach
307 to multiple testing. *J R Stat Soc B* 57: 289–300, 1995.

308 3. **Van Berkel JJBN**, **Dallinga JW**, **Moller GM**, **Godschalk RWL**, **Moonen EJ**, **Wouters EFM**, **Van**
309 **Schooten FJ**. A profile of volatile organic compounds in breath discriminates COPD patients from
310 controls. *Respir Med* 104: 557–563, 2010.

311 4. **Beurskens CJP**, **Aslami H**, **Kuipers MT**, **Horn J**, **Vroom MB**, **Van Kuilenburg ABP**, **Roelofs JJTH**,
312 **Schultz MJ**, **Juffermans NP**. Induced hypothermia is protective in a rat model of pneumococcal
313 pneumonia associated with increased adenosine triphosphate availability and turnover. *Crit Care*
314 *Med* 40: 919–926, 2012.

315 5. **Boots a W**, **Smolinska A**, **van Berkel JJBN**, **Stobberingh EE**, **Boumans MLL**, **Moonen M**, **Wouters**
316 **EFM**, **Dallinga JW**, **Schooten FJ Van**. Identification of microorganisms based on gas
317 chromatography-mass spectrometric analysis of volatile organic compounds in headspace gases. *J*
318 *Biomed Biotechnol* 8: 1752, 2014.

319 6. **Boots AW**, **van Berkel JJBN**, **Dallinga JW**, **Smolinska A**, **Wouters EF**, **van Schooten FJ**. The
320 versatile use of exhaled volatile organic compounds in human health and disease. *J Breath Res* 6:
321 027108, 2012.

322 7. **Bos L**, **Schultz M**, **Sterk P**. A simple breath sampling method in intubated and mechanically
323 ventilated critically ill patients [Online]. *Respir Physiol Neurobiol* 191: 67–74, 2014.

http://www.sciencedirect.com/science/article/pii/S1569904813003674#.

8. **Bos LDJ**, **Sterk PJ**, **Schultz MJ**. Volatile Metabolites of Pathogens: A Systematic Review. *PLoS Pathog* 9: e1003311, 2013.

9. **Bos LDJ**, **van Walree IC**, **Kolk AHJ**, **Janssen H-G**, **Sterk PJ**, **Schultz MJ**. Alterations in exhaled breath metabolite-mixtures in two rat models of lipopolysaccharide-induced lung injury. *J Appl Physiol* 115, 2013.

10. **Bos LDJ**, **Weda H**, **Wang Y**, **Knobel HH**, **Nijsen TME**, **Vink TJ**, **Zwinderman AH**, **Sterk PJ**, **Schultz MJ**. Exhaled breath metabolomics as a noninvasive diagnostic tool for acute respiratory distress syndrome. *Eur Respir J* 44: 188–197, 2014.

11. **Filipiak W**, **Beer R**, **Sponring A**, **Filipiak A**, **Ager C**, **Schiefecker A**, **Lanthaler S**, **Helbok R**, **Nagl M**, **Troppmair J**, **Amann A**. Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *J Breath Res* 9: 016004, 2015.

12. **Fowler SJ**, **Basanta-Sanchez M**, **Xu Y**, **Goodacre R**, **Dark PM**. Surveillance for lower airway pathogens in mechanically ventilated patients by metabolomic analysis of exhaled breath: a case-control study. *Thorax* (2015). doi: 10.1136/thoraxjnl-2014-206273.

13. **Grgurich PE**, **Hudcova J**, **Lei Y**, **Sarwar A**, **Craven DE**. Diagnosis of ventilator-associated pneumonia: Controversies and working toward a gold standard. *Curr Opin Infect Dis* 26: 140–150, 2013.

14. **Ibrahim B**, **Basanta M**, **Cadden P**, **Singh D**, **Douce D**, **Woodcock A**, **Fowler SJ**. Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 66: 804–809, 2011.

15. **Langford VS**, **Graves I**, **McEwan MJ**. Rapid monitoring of volatile organic compounds: A comparison between gas chromatography/mass spectrometry and selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom* 28: 10–18, 2014.

16. **Lawal O**, **Knobel H**, **Weda H**, **Bos LD**, **Nijsen TME**, **Goodacre R**, **Fowler SJ**. Volatile organic compound signature from co-culture of lung epithelial cell line with Pseudomonas aeruginosa. *Analyst* (2018). doi: 10.1039/c8an00759d.

17. **Lawal O**, **Muhamadali H**, **Ahmed WM**, **White IR**, **Nijsen TME**, **Goodacre R**, **Fowler SJ**. Headspace volatile organic compounds from bacteria implicated in ventilator-associated pneumonia analysed by TD-GC/MS. *J Breath Res* 12: 26002, 2018.

18. **Metwaly S**, **Cote A**, **Donnelly SJ**, **Banoei MM**, **Winston BW**. Evolution of ARDS biomarkers, will metabolomics be the answer? *Am J Physiol Lung Cell Mol Physiol* : 526–534, 2018.

19. **Van Oort PM**, **Povoa P**, **Schnabel R**, **Dark P**, **Artigas A**, **Bergmans D**, **Felton T**, **Coelho L**, **Schultz MJ**, **Fowler SJ**, **Bos LD**. The potential role of exhaled breath analysis in the diagnostic process of pneumonia – a systematic review. *J Breath Res* 12: 24001, 2018.

20. **Oort PMP Van**, **Bruin S De**, **Weda H**, **Knobel HH**, **Schultz MJ**, **Bos LD**. Exhaled Breath Metabolomics for the Diagnosis of Pneumonia in Intubated and Mechanically-Ventilated Intensive Care Unit (ICU) -Patients. *Int J Mol Sci* 18: 1–14, 2017.

21. **van Oort PMP**, **Nijsen T**, **Weda H**, **Knobel H**, **Dark P**, **Felton T**, **Rattray NJW**, **Lawal O**, **Ahmed W**, **Portsmouth C**, **Sterk PJ**, **Schultz MJ**, **Zakharkina T**, **Artigas A**, **Povoa P**, **Martin-Loeches I**, **Fowler SJ**, **Bos LDJ**. BreathDx – molecular analysis of exhaled breath as a diagnostic test for ventilator–associated pneumonia: protocol for a European multicentre observational study. *BMC Pulm Med* 17: 1, 2017.

22. **R Development Core Team**. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing [Online]. 2010. http://www.r-project.org.

23. **Ross BM**, **Vermeulen N**. The combined use of thermal desorption and selected ion flow tube mass spectrometry for the quantification of xylene and toluene in air. *Rapid Commun Mass Spectrom* 21: 3608–3612, 2007.

24. **Schnabel R**, **Fijten R**, **Smolinska A**, **Dallinga J**, **Boumans M-L**, **Stobberingh E**, **Boots A**, **Roekaerts P**, **Bergmans D**, **van Schooten FJ**. Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Sci Rep* 5: 17179, 2015.

25. **Schubert JK**, **Miekisch W**. Breath Analysis in Critically Ill Patients—Potential and Limitations. *Volatile Biomarkers* (2013). doi: 10.1016/B978-0-44-462613-4.00009-X.

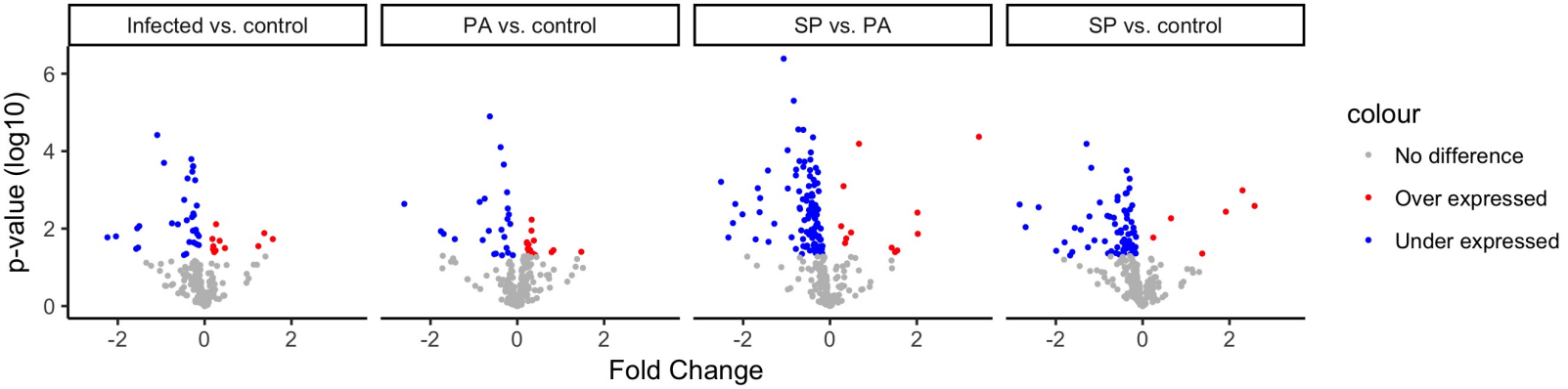
26. **Sumner LW**, **Amberg A**, **Barrett D**, **Beale MH**, **Beger R**, **Daykin CA**, **Fan TW-M**, **Fiehn O**, **Goodacre R**, **Griffin JL**, **Hankemeier T**, **Hardy N**, **Harnly J**. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Inititative (MSI). *Metabolomics* 3: 211–221, 2007.

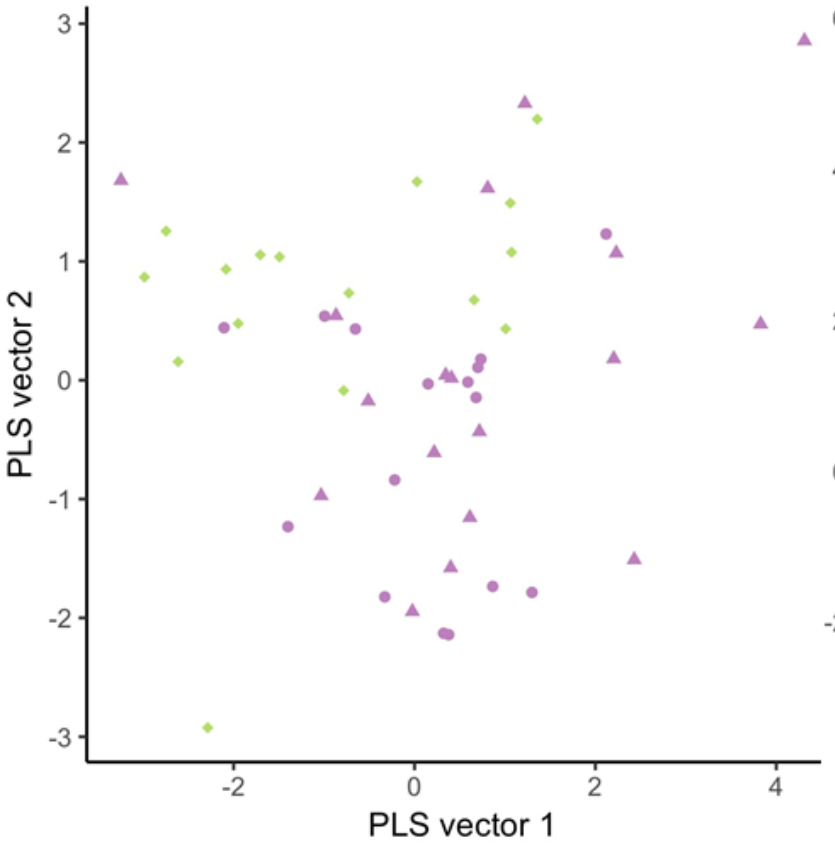
27. **Westerhuis JA**, **Hoefsloot HCJ**, **Smit S**, **Vis DJ**, **Smilde AK**, **Velzen EJJ**, **Duijnhoven JPM**, **Dorsten FA**. Assessment of PLSDA cross validation. *Metabolomics* 4: 81–89, 2008.

28. **Zhou Y**, **Chen E**, **Wu X**, **Hu Y**, **Ge H**, **Xu P**, **Zou Y**, **Jin J**, **Wang P**, **Ying K**. Rational lung tissue and animal models for rapid breath tests to determine pneumonia and pathogens. *Am J Transl Res* 9: 5116–5126, 2017.

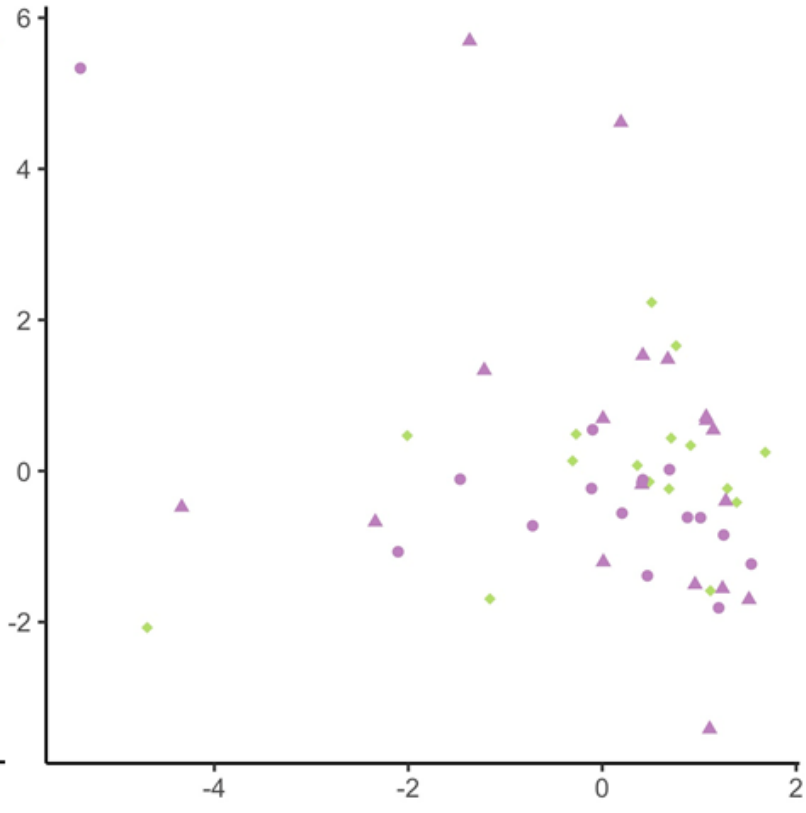
29. **Zhu J**, **Bean HD**, **Jimenez-Diaz J**, **Hill JE**. Secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting of multiple bacterial lung pathogens, a mouse model study. *J Appl Physiol* 114: 1544–1549, 2013.

30. **Zhu J**, **Jimenez-Diaz J**, **Bean HD**, **Daphtary NA**, **Aliyeva MI**, **Lundblad LKA**, **Hill JE**. Robust detection of P. aeruginosa and S. aureus acute lung infections by secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting: from initial infection to clearance. *J Breath Res* 7: 37106, 2013.

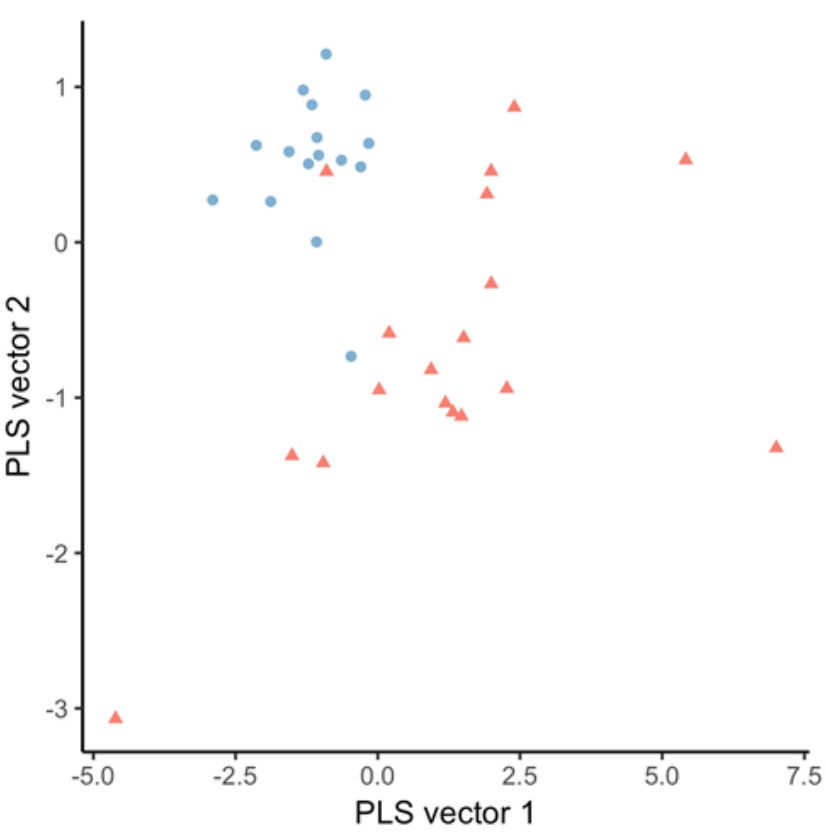




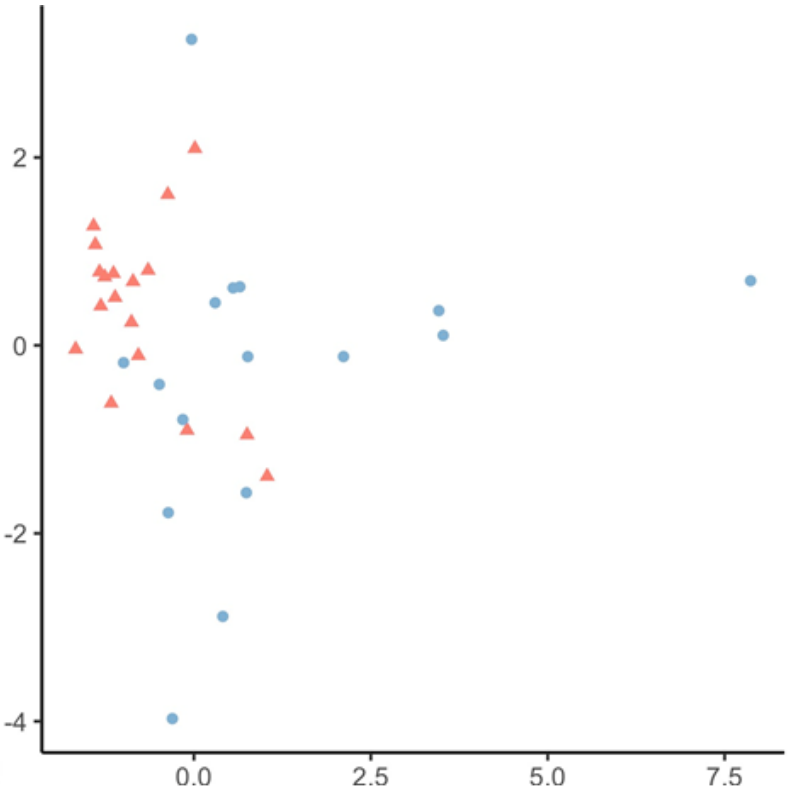
2a



2b



3a



3b

Table 1

Group comparison	VOC
Infected vs. non-infected	**Octane, 4-methyl-** **Octane, 2-5-dimethyl** **Unidentified naphthalene compound** **Unidentified cyclic compound** **Unidentified** **Unidentified branched aldehyde** Tetra chloroethylene **Unidentified cyclic compound**
SP vs. control	**Octane, 4-methyl-** **Octane, 2-5-dimethyl** **Unidentified naphthalene compound** **Unidentified** **Unidentified cyclic compound** Hexadecane Unidentified Unidentified Hexane, 2-,4-dimethyl- 2-Propanol, 1-methyloxy- Nonane, 2-methyl- Heptane, 2-,4-dimethyl **Unidentified cyclic compound** Unidentified
PA vs. control	**Unidentified branched aldehyde** 2-propenoic acid, 2-ethylhexyl ester Unidentified cyclic compound

Table 1. Identified VOCs with an adjusted *p*-value of <0.05, per comparison. In bold: identical VOCs showing overlap between the group comparisons.