Microbial volatiles as diagnostic biomarkers of bacterial lung infection in mechanically ventilated patients

Waqar M. Ahmed[1], Dominic Fenn[2,3], Iain R. White[1,4], Breanna Dixon[1], Tamara M. E. Nijsen[5], Hugo H. Knobel[6], Paul Brinkman[2], Pouline M.P. Van Oort[7], Marcus J. Schultz[8,9,10], Paul Dark[1,11], Royston Goodacre[12], Timothy Felton[1], Lieuwe D. J. Bos[2,3,9], Stephen J. Fowler[1] and the BreathDx Consortium*

1. Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, and Manchester Academic Health Science Centre and NIHR Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester, U.K.
2. Department of respiratory medicine, Amsterdam UMC location AMC, University of Amsterdam, Amsterdam, the Netherlands
3. Laboratory of Experimental Intensive Care and Anaesthesiology, Amsterdam UMC location AMC, Amsterdam, the Netherlands
4. Laboratory for Environmental and Life Science, University of Nova Gorica, Nova Gorica, Slovenia
5. Philips Research, Philips B.V., Eindhoven, the Netherlands
6. Eurofins Materials Science Netherlands BV, High Tech Campus, Eindhoven, The Netherlands
7. Department of Anaesthesiology, Amsterdam UMC location VUmc, Amsterdam, the Netherlands.
8. Intensive Care, Amsterdam UMC location AMC, Amsterdam, the Netherlands
9. Mahidol–Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand
10. Department of clinical affairs, Hamilton Medical AG, Chur, Switzerland
11. Critical Care Unit, Salford Royal NHS Foundation Trust, Northern Care Alliance NHS Group, Salford, UK
12. Centre for Metabolomics Research, Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

* BreathDx Consortium members in the acknowledgement section

Corresponding author: Stephen Fowler. Education and Research Centre, Wythenshawe Hospital, Southmoor Road, Manchester, Greater Manchester, United Kingdom M23 9LT. Email: stephen.fowler@manchester.ac.uk

Short title: Bacterial VOCs in patient breath

Abstract

Background: Early and accurate recognition of respiratory pathogens is crucial to prevent increased risk of mortality in critically ill patients. Microbial-derived volatile organic compounds (mVOCs) in exhaled breath could be used as non-invasive biomarkers of infection to support clinical diagnosis.

Methods: In this study, we investigated the diagnostic potential of *in vitro* confirmed mVOCs in the exhaled breath of patients under mechanically ventilation from the BreathDx study. Samples were analysed by thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS).

Results: Pathogens from bronchoalveolar lavage (BAL) cultures were identified in 45/89 patients and *S. aureus* was the most commonly identified pathogen (n=15). Out of 19 mVOCs detected in the *in vitro* culture headspace of four common respiratory pathogens (*Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* and *Escherichia coli*), 14 were found in exhaled breath samples. Higher concentrations of two mVOCs were found in the exhaled breath of patients infected with *S. aureus* compared to those without (3-methylbutanal p<0.01, AUROC=0.81-0.87 and 3-methylbutanoic acid p=0.01, AUROC=0.79-0.80). In addition, bacteria identified from BAL cultures which are known to metabolise tryptophan (*Escherichia coli*, *Klebsiella oxytoca* and *Haemophilus influenzae*) were grouped and found to produce higher concentrations of indole compared to breath samples with culture-negative (p=0.034) and other pathogen-positive (p=0.049) samples.

Conclusions: This study demonstrates the capability of using mVOCs to detect the presence of specific pathogen groups with potential to support clinical diagnosis. Although not all mVOCs were found in patient samples within this small pilot study, further targeted and qualitative investigation is warranted using multi-centre clinical studies.

1. Introduction

Ventilator associated pneumonia (VAP) is the most frequent nosocomial infection in critically ill patients worldwide.[1] Rapid and accurate identification of the causative pathogen is required for administration of optimal, narrow-spectrum antimicrobial therapy. Currently empirical, broad-spectrum antimicrobial therapy is recommended to avoid delays in treatment which are associated with worse clinical outcomes.[2] However, overuse of broad-spectrum antimicrobial therapy is associated with drug-induced toxicity, microbiome dysbiosis and emergence of antimicrobial resistant strains.[3–6]

A diverse range of bacteria are typically found in critically ill patients with respiratory infections (such as VAP), including gram-positive bacteria *Staphylococcus aureus, Streptococcus pneumoniae*, and gram-negative bacteria *Pseudomonas aeruginosa, Stenotrophomonas maltophilia, Haemophilus influenzae*, *Escherichia coli* and *Klebsiella pneumoniae*.[7] Current diagnostic strategies usually rely on microbiological culture of respiratory samples, which is often time-consuming (approximately three days), invasive and infrequently carried out following clinical suspicion of VAP.[8] A possible solution is to monitor molecular phenotype changes in host or bacterial metabolism.[9] There is substantial interest in the analysis of volatile organic compounds (VOCs) for the early detection of infection, primarily due to the benefits of non-invasive sample collection, adaptability of technologies, and applicability to a wide range of diseases.[10–12] Several studies have reported changes in the exhaled breath metabolome of patients with infection.[13–15] However, translational research unifying *in vitro* and in vivo studies, the important next step in biomarker validation, is lacking.

Microbial volatile organic compounds (mVOCs) are a diverse set of metabolites which originate from microbial biochemical processes and are representative of cellular signalling and metabolism.[16] Although mVOCs associated with core metabolic pathways are consistently produced, the volatilome of a microbe differs between various taxa and are highly influenced by the surrounding environment (nutrients, pH and CO_2/O_2 levels). This makes it difficult to pinpoint the causative pathogen in a study designed to focus solely on a type of infection, rather than commensal resident microbes,

particularly in VAP where several causative pathogens are routinely identified. Instead, volatile metabolites are related to infection pathogenesis as a whole, which may include physiological and host response factors, as demonstrated by previous untargeted studies which advocate a high sensitivity or "rule-out" test method.[17–19] Alternatively, a limited number of studies have focused solely on mVOCs, for VAP and Invasive Aspergillosis in critically ill patients.[20–22]

In this study, the diagnostic potential of measuring mVOCs were evaluated. Culture headspace experiments of bacterial species associated with VAP informed which mVOCs to target in patient breath samples. Compounds detected in breath were then compared in *culture-negative* (< 10^4 colony forming units [CFU] mL^{-1}) and *other pathogen-positive* (all other samples with pathogens identified in BAL culture apart from the pathogen of interest for statistical comparison) groups. The outcome of this study demonstrated the feasibility of mVOC biomarkers to detect the presence of a known pathogen and potentially guide early and appropriate antimicrobial treatment.

2. Methods

2.1. Headspace sample collection

The reference strains *S. aureus* ATCC 29213, *P. aeruginosa* PAO1, *K. pneumoniae* ATCC 13887, and *E. coli* ATCC 25922 were used in this study. All strains were cultured in nutrient broth (Sigma Aldrich, St. Louis, US). For headspace samples, axenic cultures were standardised to OD_{600nm} 0.01 to a total volume of 2 mL (~ $1x10^6$ to $1x10^7$ CFU mL^{-1}) were crimp-sealed in 20 mL glass vials with a PTFE septum and incubated for 24 h at 37 °C under shaking conditions. Two headspace gas sampling methods were used to capture mVOCs, as described previously.[23] Further details are provided in the supplementary file.

2.2. Patient sample collection

Exhaled breath sample data from the BreathDx study were used (UKCRN ID 19086, May 2015). During this study, 1200 mL exhaled breath from ventilator circuit tubing was sampled into TenaxGR sorbent tubes (flow rate of 200 mL min^{-1}) as previously described.[24] Eighty-nine patients from the

Intensive Care Unit (ICU) departments of four university hospitals, Salford Royal NHS Foundation Trust (*n* = 29), Academic Medical Centre Amsterdam (*n* = 42), Central Manchester NHS Foundation Trust (*n* = 16) and Wythenshawe Hospital NHS Foundation Trust (*n* = 2), were recruited over a two-year period from February 2016 to February 2018. Semi-quantitative bronchoalveolar lavage (BAL) cultures were used as the standard reference for pathogen identification with a cut-off of ≥ 10^4 CFU mL^{-1} to define a positive culture. Exhaled breath and BAL were sampled within 24 h of clinical suspicion of VAP and before antimicrobial treatment.

2.3. Analysis by thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS)

Culture headspace and breath samples were analysed by TD-GC-MS using the method described previously.[24] For further details see the Supplementary file.

2.4. Data pre-processing

Initial tests were performed on full-scan data that were previously pre-processed for the preliminary Breathdx untargeted analysis described in van Oort et al. (2022). This data set, comprising 1247 MS features, was tested for differences between pathogen-positive and pathogen-negative groups. A batch search method was then created in Masshunter Quantitative analysis software (B.07.00, Agilent technologies, Cheadle, UK) based on chromatograms from culture headspace samples. Peaks were selected if they were present in at least one pathogen across all culture replicates. Features that were present in culture *versus* media control samples were not assessed. The batch method was then used to determine breath analytes. Mass spectra and retention indices of external standards were used to support identification (i.e. selection of appropriate compound quantifier and qualifier ions, and retention time windows), adhering to the metabolomics standards initiative.[25,26] Missing values were allocated based on the determination of qualifier ions. An *n* x *p* data matrix was created for subsequent data treatment and statistical analyses. Estimated VOC concentrations were determined from a standard curve created using a calibration standards loading rig (Markes International, Bridgend, UK).

2.5. Statistical analysis

Statistical analyses were carried out in GraphPad Prism (v. 9.1) and R (v 4.1.1). All sample intensities were normalised using a spiked internal standard (4-bromofluorobenzene). Volcano plots were made from the untargeted dataset and significant features were listed. Peak intensities from the *in vitro* culture study were mean centred and hierarchical cluster analysis was carried out on all in the R package *pheatmap* generating a heatmap for visual comparisons. For breath sample data, missing values were imputed with a random value between zero and the limit of detection (3 x σ of background samples) for each compound. Intraclass correlation coefficients were calculated between repeat breath samples from the same patient.

The non-parametric Kruskal-Wallis test was applied to screen for statistical differences between groups. This was followed by a *post hoc* Dunn's multiple comparison test between two independent groups. Individual p values were reported (α = 0.05) and no correction was made for multiple comparisons. For the *in vitro* culture study, differences between sampling methods were assessed. For breath samples, differences between pathogen-positive, other pathogen positive, and culture-negative groups were assessed. The untargeted dataset was analysed using the univariate Mann-Whitney U test between pathogen-positive and pathogen-negative groups where statistical significance and fold change were illustrated using volcano plots. To explore the diagnostic performance of any potential VOC markers, the area under the receiver operating characteristic (AUROC) curve was calculated.

3. Results

3.1. Patient characteristics

Patient characteristics, stratified for positive and negative BAL cultures, are summarised in Table 1. Differences were found between admission type (medical compared to non-medical) and the Clinical Pulmonary Infection Score (CPIS) between culture-positive and culture-negative patients.

3.2. Microbial VOCs from culture headspace

A total of 19 mVOCs were detected from the headspace of axenic cultures. Compounds detected in control samples including culture media, sampling equipment, and the instrument background, were not included. Compounds which could not be confidently detected due to co-elution or were outside the pre-defined criteria, were not analysed further. The heatmap in Figure 1 highlights the differences in mVOC abundance across microbes and sampling methods.

Based on volatile profiles alone, all sample replicates clustered by microorganism, regardless of the headspace sampling method used. Acetone, dimethyl disulfide and dimethyl trisulfide were found in the headspace of all bacteria. No other mVOCs were found to be consistently produced in all microbes investigated in this study. Some mVOCs were only observed in one microbial species compared to others. These included 1-undecene, methyl thiocyanate, dimethyl sulfide, and 2-aminoacetophenone from *P. aeruginosa* cultures; ethyl acetate, 2-heptanone, and 2-nonanone from *K. pneumoniae* cultures; and benzaldehyde and indole from *E. coli* cultures. Several mVOCs were found across two or more species such as 3-methyl-1-butanol, 3-methylbutanal, and 3-methylbutanoic acid which were higher in *K. pneumoniae* and *S. aureus* but were not detected in *P. aeruginosa* and *E. coli*, with the exception of 3-methylbutanal for *E. coli*.

Differences between the two sampling methods employed were also observed. Methyl thiocyanate and dimethyl sulfide, produced by *P. aeruginosa*, were detected using active sampling only. Conversely, compounds detected solely by passive sampling included 2-aminoacetophenone from *P. aeruginosa* ($p = 0.036$), and 3-methylbutanoic acid ($p = 0.007$) and furfuryl formate ($p < 0.001$) from *S. aureus*. A significant increase in dimethyl disulfide from *P. aeruginosa* was demonstrated (*K. pneumoniae* $p < 0.001$, *E. coli* $p = 0.003$, *S. aureus* $p < 0.001$) using the same passive method.

3.3. Microbial VOCs in patient exhaled breath samples from the ventilator circuit

Initial tests with full scan data previously analysed in the preliminary BreathDx study (van Oort et al. 2022) revealed significant differences between pathogen-positive and pathogen-negative cultures, as shown in the volcano plots (see Supplementary Figure S1). Several significant features ($p < 0.05$) were found for *S. aureus*, *P. aeruginosa*, Klebsiella species, and Indole-producing pathogens when compared to culture-negative samples (n = 24, n = 21, n=10, and n = 143, respectively).

To explore mVOCs produced by pathogens, a search of mVOCs detected in culture headspace was carried out on patient breath samples. Five mVOCs were not confidently measured in breath (e.g., due to co-elution or because they were outside pre-defined criteria) and were therefore excluded from further statistical analysis, namely 2-propanol, methyl thiocyanate, furfuryl formate, 2,5-dimethylpyrazine, and dimethyl trisulfide. Significant changes between pathogen-positive, other pathogen-positive, and culture-negative samples were assessed for mVOCs detected in breath samples (Table 2). Breath sample reproducibility was assessed by calculating ICC between sequential samples from the same patient and ranged 0.49 to 0.98 for mVOCs detected in patients' breath (Supplementary Table S1).

Acetone and benzaldehyde were detected in all breath samples. Acetone was also significantly lower in *P. aeruginosa*-positive compared to other pathogen-positive samples ($p = 0.007$). A lower abundance of 3-methylbutanal was also observed for *P. aeruginosa* compared to other pathogen-positive samples ($p = 0.012$). In samples with positive culture identification for *S. aureus*, both 3-methylbutanal and 3-methylbutanoic acid were significantly higher compared to culture-negative samples ($p < 0.001$, $p = 0.001$, respectively). When compared to other pathogen-positive samples, 3-methylbutanal and 3-methylbutanoic acid were also significantly higher in *S. aureus*-positive samples ($p < 0.001$, $p = 0.001$, respectively). Importantly, no significant differences were found for both compounds between other pathogen-positive samples or between culture positive vs culture negative samples.

Diagnostic performance was assessed for *S. aureus*-positive samples against culture-negative or other pathogen-positive groups which resulted in an AUROC of 0.87 (95% CI 0.75-0.99, p <0.001) and 0.81 (95% CI 0.66-0.96, p < 0.001) for 3-methylbutanal, respectively. For 3-methylbutanoic acid, *S. aureus*-positive versus culture-negative and other pathogen-positive resulted in an AUROC of 0.79 (95% CI 0.65-0.93, p = 0.001) and 0.80 (95% CI 0.64-0.95, p = 0.002), respectively. Pearson's correlation coefficient showed a weak positive correlation (r = 0.21, p = 0.045) between the two compounds. Results for 3-methylbutanal and 3-methylbutanoic acid are shown in Figure 2. Due to the low number of *E. coli*-positive (n=2) and *K. pneumoniae*-positive (n=2) cultures, these samples were excluded from further statistical analysis. Although not confirmed using culture headspace in this study, grouping breath samples with bacteria which can metabolise tryptophan (n=9) *Klebsiella oxytoca*, *Haemophilus influenza* and *E. coli* displayed higher abundance of indole in breath samples (against other pathogen-positive p = 0.049 and culture-negative p = 0.034). No significant differences were found for *Klebsiella* species alone (n=5) when compared to other pathogen-positive or culture-negative samples.

4. Discussion

Exhaled breath biomarker discovery for detecting microbial infection in critically ill patients remains a challenge. This study investigates the value of targeting mVOCs as diagnostic markers of bacterial infection and we have presented evidence for non-invasive detection of *S. aureus* infection and indole-producing pathogens from in a multi-centre ICU study.

We screened patient breath sample data in an untargeted approach which revealed several significantly different mass spectral features. These features are indicative of bacterial infection including host response and not mVOCs produced by a specific pathogen and group of pathogens. We then developed a target list of mVOCs based on culture headspace data and searched for them in exhaled breath samples collected non-invasively from the ventilator circuit. The most significant result was higher abundance of 3-methylbutanal (syn. isovaleraldehyde) and 3-methylbutanoic acid

(syn. isovaleric acid) produced by *S. aureus in vitro* and in exhaled breath from patients with confirmed *S. aureus* infection. Importantly, production of both compounds was higher compared to both other pathogen-positive and culture-negative samples. Previous studies have also detected 3-methylbutanal and/or 3-methylbutanoic acid associated with *S. aureus* from culture headspace[27–35] and ventilator exhaled breath,[21] irrespective of culture conditions, strain type and sampling methods. As both compounds were found to be correlated with each other, it is likely they share the same metabolic pathway. Branched chain fatty acids (e.g. 3-methylbutanoic acid), and their precursors (e.g. 3-methylbutanal) from leucine metabolism are important for maintaining membrane fluidity in *S. aureus*.[36–38] Indole was also detected in higher abundance in bacteria known to use the enzyme tryptophanase to catabolise the essential amino acid tryptophan.[39] Previous studies have detected indole in the headspace of *H. influenzae*[40] and *K. oxytoca*[41] and we confirmed indole production from *E. coli* cultures in this study. This highlights the importance of linking mVOCs with their metabolic pathways which may not be species-specific and instead shared between bacterial classifications and intrinsic biochemical processes.

The number of pathogens identified from BAL culture limited statistical comparisons of several bacterial species (e.g. for *E. coli*, present in culture from only two patients). In addition, previously reported microbial volatiles from literature were not included in this analysis due to methodological differences. Culture headspace samples were analysed using the same untargeted TD-GC-MS method as in BreathDx, limiting an extensive analysis of all mVOCs. Along with developing an analytical method around known mVOCs, breath sampling methods and influence of bacterial growth phase should be carefully considered in future studies depending on the target analytes. In addition, it is important to state that although commonly reported mVOCs were identified under laboratory growth conditions, they may not represent a ventilated human lung environment exposed to environmental and drug-induced stresses.[32,42,43] Additionally, clinical isolates and antimicrobial resistant strains have also shown variation in their volatilomes.[44–46] More investigative work is required in the form of lung infection models and metabolic flux studies.[32,47–49] Some mVOCs

identified from bacteria (e.g. ethyl acetate) are also identified in the headspace of fungi.[23,50–52] Other compounds such as benzaldehyde are known environmental or sorbent artefacts. Acetone is commonly found in almost all breath samples from healthy individuals. Such compounds are challenging to validate as pathogen or disease biomarkers.

Exhaled breath sampling from patients on mechanical ventilation is not without its challenges. It has previously been shown that different ventilator systems influence background VOC profiles differently and vary in the extent to which they dilute VOC recovery.[53] In the BreathDx study, potential confounders including storage time, analysis date, and ventilation parameters were not found to influence breath volatile profiles between patients with or without infection. Samples obtained from one of the sites in the study could be separated by principal component analysis (second dimension) and this may be due to site-specific practices, such as preference for anaesthetic medication, which could not be controlled.[19]

The effect of medication on breath profiles has been observed previously[54] and although patients received anti-infective treatment after breath analysis, effect of other medication across the patient cohort is a potential confounder. For example, anaesthetic drugs and their metabolites were detected in chromatograms where propofol administered < 24h had significantly increased peak area ($p=0.009$) compared to patients administered propofol > 24h (see Supplementary Figure S2). Progress has been made in the standardisation of breath VOC analysis by preconcentration, however there are several factors limiting current clinical use such as sample storage and water retention.[12,55–58] The value of VOCs as diagnostic biomarkers is heavily influenced by the value of the clinical information collected. Due to the highly sensitive nature of VOC analysis, differences between hospital sites, clinical workflows and diagnostic models, ventilator settings, or preferred anti-infective treatment regime, will be reflected in the metabolic phenotype. Although breath was collected on suspicion of VAP and before antibiotic treatment, critically ill patients would already have multiple co-morbidities and administered several medications, for example anaesthetic and

anti-inflammatories, which may likely change the composition of microbiomes.[59] It is widely regarded that the lungs are not a sterile environment,[60] therefore applying a conservative CFU cut-off value for BAL culture may increase the likelihood of false-positive identifications and impact diagnostic accuracy of measuring mVOCs. Future studies should therefore consider measuring microbiome diversity and abundance and how they correlate to exhaled breath VOCs for those pathogens below the clinical CFU threshold.

Microbial volatiles in exhaled breath may be interpreted as a "gas lavage" of the lungs and representative of systemic infection compared to localised and invasive wash. We have demonstrated using a translational approach that mVOCs have diagnostic utility in detecting pathogens using exhaled breath samples. Further validation studies are required as not all mVOCs were found in breath samples in our study. A targeted assay of mVOCs may be clinical useful in reducing healthcare-associated infection and minimising use unnecessary use of antimicrobials.

NOTES

Contributions SJF, TMEN, LDJB, MJS, HHK, PD, RG, WA and TF were involved in study design and methodology of the BreathDx study. WA, IRW, BD, PB, DF, PMPVO and LDJB were involved in breath and headspace sampling and analysis. WA wrote the manuscript.

* BreathDx Consortium members in the acknowledgement section: Waqar M. Ahmed, Antonio Artigas Raventos, 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, Craig Johnson, Hugo Knobel, Oluwasola Lawal, Jan-Hendrik Leopold, Ignacio Martin-Loeches, Tamara M. E. Nijsen, Pouline M. P. van Oort, Pedro Povoa, 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.

Funding

The project was funded by the European Union FP7 Marie Curie Actions, under the Industry-Academia Partnerships and Pathways (IAPP) programme (MC-IAPP BreathDx 611951), and supported by the NIHR Manchester Biochemical Research Centre and the NIHR Manchester Clinical

Research Facility. IW is support by the Slovenian research agency (ARRS) through program P1-0034: Analytics and Chemical Characterization of Materials and Processes. BD reports support from Biotechnology and Biological Sciences Research Council [Grant No BB/T008725/1] (PhD studentship, funded to institution).

Potential Conflicts of Interest:

PD is Deputy Medical Director, NIHR National Research Network (.5 FTE payment to institution), and NIHR Senior Investigator (£20,000 per annum to my employing institution). BD reports travel accommodation from Markes International for IABR conference 2022. LB reports consulting fees from Scailyte and Santhera (paid to institution), and participation on advisory board for Sobi, Excastat, Pfizer, and Astra Zeneca. PB reports grants for development of chemical sensor driven technology from Amsterdam UMC - Innovatie Impuls 2020, Vertex - Vertex Innovation Award 2022, Stichting Astma Bestrijding (SAB), and Boehringer Ingelheim; Public-Private Partnership grant from Eurostars; and grant for Disaster-Resilient Society 2021 (HORIZON-CL3-2021-DRS-01) from Horizon Europe Framework Programme (HORIZON). All authors approved the final version of the manuscript. All other authors declare no conflicts of interest.

References

1. Timsit, J.-F., Esaied, W., Neuville, M., Bouadma, L. & Mourvillier, B. Update on ventilator-associated pneumonia. *F1000Research* **6**, 2061 (2017).

2. Torres, A. *et al.* International ERS/ESICM/ESCMID/ALAT guidelines for the management of hospital-acquired pneumonia and ventilator-associated pneumonia. *Eur. Respir. J.* **50**, (2017).

3. Wunderink, R. G. *et al.* Antibiotic stewardship in the intensive care unit : An official American thoracic society workshop report in collaboration with the aacn, chest, cdc, and sccm. *Ann. Am. Thorac. Soc.* **17**, 531–540 (2020).

4. Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **6736**, 02724 (2022).

5. Hagan, T. *et al.* Antibiotics-Driven Gut Microbiome Perturbation Alters Immunity to Vaccines in Humans. *Cell* **178**, 1313-1328.e13 (2019).

6. Tamma, P. D., Avdic, E., Li, D. X., Dzintars, K. & Cosgrove, S. E. Association of adverse events with antibiotic use in hospitalized patients. *JAMA Intern. Med.* **177**, 1308–1315 (2017).

7. Koulenti, D., Tsigou, E. & Rello, J. Nosocomial pneumonia in 27 ICUs in Europe: perspectives from the EU-VAP/CAP study. *Eur. J. Clin. Microbiol. Infect. Dis.* **36**, 1999–2006 (2017).

8. Rea-Neto, A. *et al.* Diagnosis of ventilator-associated pneumonia: A systematic review of the literature. *Crit. Care* **12**, 1–14 (2008).

9. Bos, L. D. J., Sterk, P. J. & Schultz, M. J. Volatile Metabolites of Pathogens: A Systematic Review. *PLoS Pathog.* **9**, e1003311 (2013).

10. de Lacy Costello, B. *et al.* A review of the volatiles from the healthy human body. *J. Breath Res.* **8**, 014001 (2014).

11. Bos, L. D., Sterk, P. J. & Fowler, S. J. Breathomics in asthma and COPD. *J. Allergy Clin. Immunol.* (2016) doi:10.1016/j.jaci.2016.08.004.

12. Horváth, I. *et al.* A European Respiratory Society technical standard : exhaled biomarkers in lung disease. *Eur Respir J* **49**, 1–26 (2017).

13. Ahmed, W. M., Lawal, O., Nijsen, T. M., Goodacre, R. & Fowler, S. J. Exhaled Volatile Organic Compounds of Infection: A Systematic Review. *ACS Infect. Dis.* acsinfecdis.7b00088 (2017) doi:10.1021/acsinfecdis.7b00088.

14. Lawal, O., Ahmed, W. M., Nijsen, T. M. E., Goodacre, R. & Fowler, S. J. Exhaled breath analysis: a review of 'breath-taking' methods for off-line analysis. *Metabolomics* **13**, 1–16 (2017).

15. Van Oort, P. M. *et al.* The potential role of exhaled breath analysis in the diagnostic process of pneumonia – a systematic review. *J. Breath Res.* **12**, 24001 (2018).

16. Schmidt, R., Cordovez, V., De Boer, W., Raaijmakers, J. & Garbeva, P. Volatile affairs in microbial interactions. *ISME J.* **9**, 2329–2335 (2015).

17. Fowler, S. J., Basanta-Sanchez, M., Xu, Y., Goodacre, R. & Dark, P. M. Surveillance for lower airway pathogens in mechanically ventilated patients by metabolomic analysis of exhaled breath: a case-control study. *Thorax* **70**, 320–325 (2015).

18. Schnabel, R. *et al.* Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia. *Sci. Rep.* **5**, 1–10 (2015).

19. Van Oort, P. M. P. *et al.* Untargeted Molecular Analysis of Exhaled Breath as a Diagnostic Test for Ventilator-Associated Lower Respiratory Tract Infections (BreathDx). *Thorax* 1–3 (2021) doi:10.1136/thoraxjnl-2021-217362.

20. Koo, S. *et al.* A Breath Fungal Secondary Metabolite Signature to Diagnose Invasive Aspergillosis. *Clin. Infect. Dis.* **59**, 1733–1740 (2014).

21. Filipiak, W. *et al.* 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).

22. Gao, J. *et al.* Breath analysis for noninvasively differentiating *Acinetobacter baumannii* ventilator-associated pneumonia from its respiratory tract colonization of ventilated patients. *J. Breath Res.* **10**, 027102 (2016).

23. Ahmed, W. M. *et al.* Development of an adaptable headspace sampling method for metabolic profiling of the fungal volatome. *Analyst* (2018) doi:10.1039/C8AN00841H.

24. van Oort, P. M. P. *et al.* 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**, 1–8 (2017).

25. Goodacre, R. *et al.* Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* **3**, 231–241 (2007).

26. Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).

27. Chen, J., Tang, J., Shi, H., Tang, C. & Zhang, R. Characteristics of volatile organic compounds produced from five pathogenic bacteria by headspace-solid phase micro-extraction/gas chromatography-mass spectrometry. *J. Basic Microbiol.* **57**, 228–237 (2017).

28. Jünger, M. *et al.* Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. *Appl. Microbiol. Biotechnol.* **93**, 2603–2614 (2012).

29. Preti, G. *et al.* Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: Analysis by solid-phase microextraction and gas chromatography-mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **877**, 2011–2018 (2009).

30. Zechman, J. M., Aldinger, S. & Labows, J. N. Characterization of pathogenic bacteria by automated headspace concentration-gas chromatography. *J. Chromatogr. B. Biomed. Sci. App.* **377**, 49–57 (1986).

31. Fitzgerald, S., Duffy, E., Holland, L. & Morrin, A. Multi-strain volatile profiling of pathogenic and commensal cutaneous bacteria. *Sci. Rep.* **10**, 1–12 (2020).

32. Lawal, O. *et al.* Headspace volatile organic compounds from bacteria implicated in ventilator-associated pneumonia analysed by TD-GC/MS. *J. Breath Res.* **12**, (2018).

33. Tait, E., Perry, J. D., Stanforth, S. P. & Dean, J. R. Identification of volatile organic compounds produced by bacteria using HS-SPME-GC-MS. *J. Chromatogr. Sci.* **52**, 363–373 (2014).

34. Jia, B. *et al.* Distinguishing Methicillin-Resistant and Sensitive Staphylococcus aureus Using Volatile Headspace Metabolites. *IEEE Sens. J.* **10**, 71–75 (2010).

35. Filipiak, W. *et al.* Molecular analysis of volatile metabolites released specifically by staphylococcus aureus and pseudomonas aeruginosa. *BMC Microbiol.* **12**, 113 (2012).

36. Frank, M. W., Whaley, S. G. & Rock, C. O. Branched-chain amino acid metabolism controls membrane phospholipid structure in Staphylococcus aureus. *J. Biol. Chem.* **297**, 1–14 (2021).

37. Parsons, J. B. & Rock, C. O. Bacterial lipids: Metabolism and membrane homeostasis. *Prog. Lipid Res.* **52**, 249–276 (2013).

38. Chen, X., Ping Teoh, W., Stock, M. R., Resko, Z. J. & Alonzo, F. Branched chain fatty acid synthesis drives tissue-specific innate immune response and infection dynamics of Staphylococcus aureus. *PLoS Pathog.* **17**, 1–22 (2021).

39. Roager, H. M. & Licht, T. R. Microbial tryptophan catabolites in health and disease. *Nat. Commun.* **9**, 3294 (2018).

40. Preti, G. *et al.* Volatile compounds characteristic of sinus-related bacteria and infected sinus mucus: Analysis by solid-phase microextraction and gas chromatography-mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **877**, 2011–2018 (2009).

41. Jünger, M. *et al.* Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. *Appl. Microbiol. Biotechnol.* **93**, 2603–2614 (2012).

42. Zhu, J., Bean, H. D., Wargo, M. J., Leclair, L. W. & Hill, J. E. Detecting bacterial lung infections: in vivo evaluation of in vitro volatile fingerprints. *J. Breath Res.* **7**, 016003 (2013).

43. Jenkins, C. L. & Bean, H. D. Influence of media on the differentiation of Staphylococcus spp. By volatile compounds. *J. Breath Res.* **14**, (2020).

44. Bean, H. D., Rees, C. A. & Hill, J. E. Comparative analysis of the volatile metabolomes of Pseudomonas aeruginosa clinical isolates. *J. Breath Res.* **10**, (2016).

45. Smart, A. *et al.* Sniffing out resistance – Rapid identification of urinary tract infection-causing bacteria and their antibiotic susceptibility using volatile metabolite profiles. *J. Pharm. Biomed. Anal.* **167**, 59–65 (2019).

46. Rees, C. A. *et al.* Detection of high-risk carbapenem-resistant Klebsiella pneumoniae and Enterobacter cloacae isolates using volatile molecular profiles. *Sci. Rep.* **8**, 1–13 (2018).

47. Schivo, M. *et al.* Volatile emanations from in vitro airway cells infected with human rhinovirus. *J. Breath Res.* **8**, (2014).

48. Traxler, S. *et al.* Volatile scents of influenza A and S. pyogenes (co-)infected cells. *Sci. Rep.* **9**, 1–12 (2019).

49. Lawal, O. *et al.* Volatile organic compound signature from co- culture of lung epithelial cell line with Pseudomonas aeruginosa. *Analyst* (2018) doi:10.1039/c8an00759d.

50. Hertel, M. *et al.* Identification of signature volatiles to discriminate Candida albicans, glabrata, krusei and tropicalis using gas chromatography and mass spectrometry. *Mycoses* **59**, 117–126 (2015).

51. Thorn, R. M. S., Reynolds, D. M. & Greenman, J. Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro. *J. Microbiol. Methods* **84**, 258–264 (2011).

52. Filipiak, W. *et al.* Characterization of volatile metabolites taken up by or released from Streptococcus pneumoniae and Haemophilus influenzae by using GC-MS. *Microbiology* **158**, 3044–3053 (2012).

53. Van Oort, P. M. P. *et al.* Detection and quantification of exhaled volatile organic compounds in mechanically ventilated patients-comparison of two sampling methods. *Analyst* **146**, 222–231 (2021).

54. Kovacs, D., Bikov, A., Losonczy, G., Murakozy, G. & Horvath, I. Follow up of lung transplant recipients using an electronic nose. *J. Breath Res.* **7**, 017117 (2013).

55. Wilkinson, M., White, I. R., Goodacre, R., Nijsen, T. & Fowler, S. J. Effects of high relative humidity and dry purging on VOCs obtained during breath sampling on common sorbent tubes. *J. Breath Res.* **14**, (2020).

56. Wilkinson, M. *et al.* The peppermint breath test: A benchmarking protocol for breath sampling and analysis using GC-MS. *J. Breath Res.* **15**, (2021).

57. Kang, S. & Paul Thomas, C. L. How long may a breath sample be stored for at −80 °C? A study of the stability of volatile organic compounds trapped onto a mixed Tenax:Carbograph trap adsorbent bed from exhaled breath. *J. Breath Res.* **10**, 026011 (2016).

58. Ahmed, W. *et al.* Methodological considerations for large-scale breath analysis studies: lessons from the U-BIOPRED severe asthma project. *J. Breath Res.* **13**, 016001 (2019).

59. Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–628 (2018).

60. Dickson, R. P. & Huffnagle, G. B. The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease. *PLoS Pathog.* **11**, 1–5 (2015).

Table 1: Patient clinical characteristics

		Positive BAL culture *n*= 45 (50.6 %)	**Negative BAL culture** *n*= 44 (49.4 %)	*p*-value
Age in years (median, IQR)		58.5 (39.25-68.75)	59 (46-66)	0.788 [b]
Gender (*n*, %)	Male	27 (60)	30 (68.2)	0.509 [c]
	Female	18 (40)	14 (31.8)	
BMI (median, IQR)		24.4 (20.9-31.0)	26.2 (22.9-32.0)	0.290 [b]
Admission type (*n*, %) [a]	Medical	15 (33.3)	27 (61.4)	0.011 [c]
	Planned surgical	16 (35.6)	6 (13.6)	0.364 [c, d]
	Emergency surgical	14 (31.1)	10 (22.7)	
Admission characteristics (*n*, %)	Trauma	18 (40)	9 (20.5)	0.065 [c]
	Neurosurgery	15 (33.3)	8 (18.2)	0.146 [c]
Clinical scores (median, IQR)	SOFA lung	4 (2-4)	3 (2-4)	0.118 [b]
	APACHEII	17 (10-23)	20 (14.3-23.8)	0.389 [b]
	CPIS	7 (5.5-7)	5 (4-6)	< 0.001 [b]
ICU LOS in days (median, IQR)		21 (15-32)	21 (13-33)	0.926 [b]
ICU mortality (*n*, %)		7 (15.6)	11 (25)	0.302 [c]
Pathogen isolated (*n*, %) [e]	*S. aureus*	15 (33.3)		
	P. aeruginosa	8 (17.8)		
	E. coli	2 (4.4)		
	K. pneumoniae	2 (4.4)		
	Other	24 (53.3)		
	(indole producing [f])	9 (20.0)		

BAL = Bronchoalveolar lavage, IQR = interquartile range, BMI = body mass index, ICU = Intensive care unit, LOS = Length of stay. [a] one patient with no information in the negative culture group, [b] Mann-Whitney-U test, [c] Fisher's exact test, [d] comparison between planned and emergency surgical admission type, [e] Potentially >1 cultured pathogen per patient, [f] pathogens which are known to produce indole *E. coli* (n=2), *H. influenzae* (n=5), and *K. oxytoca* (n=2).

Table 2. Microbial volatiles detected in patient samples with a comparison of pathogen-positive samples

Compound name [+]	CAS [*]	Synonym	*m/z*		Retention index [#]	Detected in *n* breath samples	Relative concentration and *p* value	
			Quantifier ion	Qualifier ion(s)			Pathogen +ve / other pathogen +ve	Pathogen +ve / culture -ve
acetone	67-64-1	propanone	58	-	505	89	↓ 0.007 (*P. aeruginosa*)	-
dimethyl sulfide	75-18-3	dimethyl thioether	47	62	566	70	-	-
ethyl acetate	141-78-6	acetic acid, ethyl ester	43	88	607	73	-	-
3-methylbutanal	590-86-3	isovaleraldehyde	58	71, 86	647	59	↓ 0.012 (*P. aeruginosa*) ↑ < 0.001 (*S. aureus*)	↑ < 0.001 (*S. aureus*)
3-methyl-1-butanol	123-51-3	isopentyl alcohol	55	70	730	59	-	-
dimethyl disulfide	624-92-0	(methyldithio)methane	94	45, 79	769	57	-	-
3-methylbutanoic acid	503-74-2	isovaleric acid	60	87	858	74	↑ 0.001 (*S. aureus*)	↑ 0.001 (*S. aureus*)
2-heptanone	110-43-0	heptan-2-one	58	71	896	73	-	-
benzaldehyde	100-52-7	benzoic aldehyde	77	106	969	89	-	-
2-nonanone	821-55-6	nonan-2-one	58	71, 142	1088	82	-	-
1-undecene	821-95-4	undec-1-ene	55	70, 83, 97, 154	1101	26	-	-
2-phenylethanol	60-12-8	phenylethyl alcohol	91	91.9, 122	1122	71	-	-
indole	120-72-9	benzopyrrole	90	117	1278	84	-	-
2-aminoacetophenone	551-93-9	ethanone, 1-(2-aminophenyl)-	120	92, 135	1312	37	-	-

+ mVOCs were identified using the mass spectra and retention times of analogous external standards run on the same TD-GC-MS (i.e. MSI level 1). # Calculated Kováts Retention index using an alkane ladder (C_5-C_{15}) on a standard non-polar column. * Chemical Abstracts Service (CAS) registry number.

FIGURE LEGENDS

Figure 1: *Heatmap ordered by hierarchical cluster analysis for mVOCs and individual samples produced by different pathogens and sampling methods (n=5 replicates for Active sampling, n=3 replicates for Passive sampling). The colour scale (blue to red) represents VOC peak intensity z-scores from axenic culture headspace. Sampling methods and microorganisms are colour coded above the heatmap, and details are provided within the figure.*

Figure 2: *Scatter plots of estimated VOC concentrations and ROC curves of 3-methylbutanal (A and B) and 3-methylbutanoic acid (C and D). Scatter plots are annotated with the group median (black line), estimated limit of detection (dotted line) and p-values. Comparisons were made between S. aureus-positive and other pathogen-positive or culture-negative samples.*

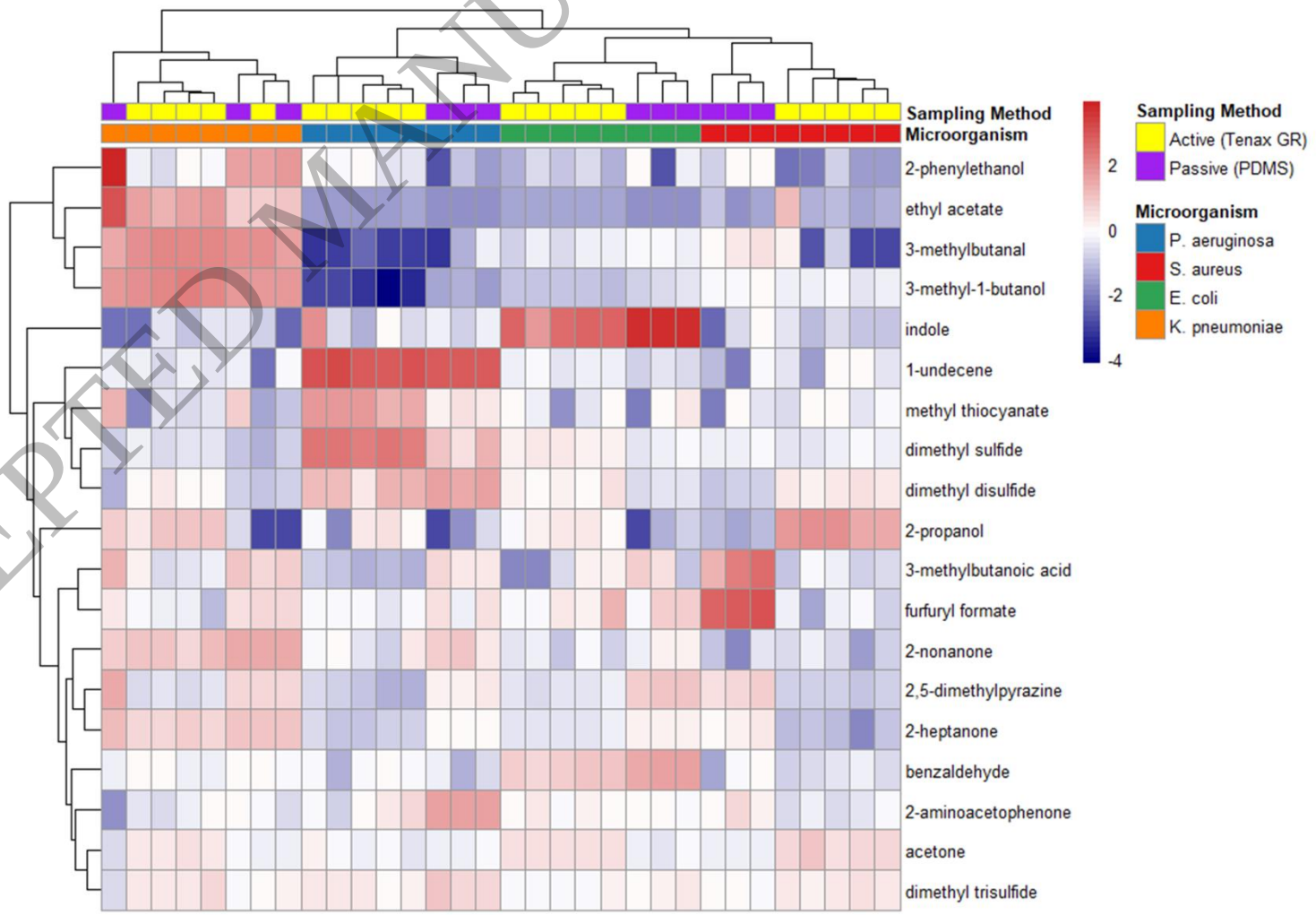

Figure 1
339x190 mm (x DPI)

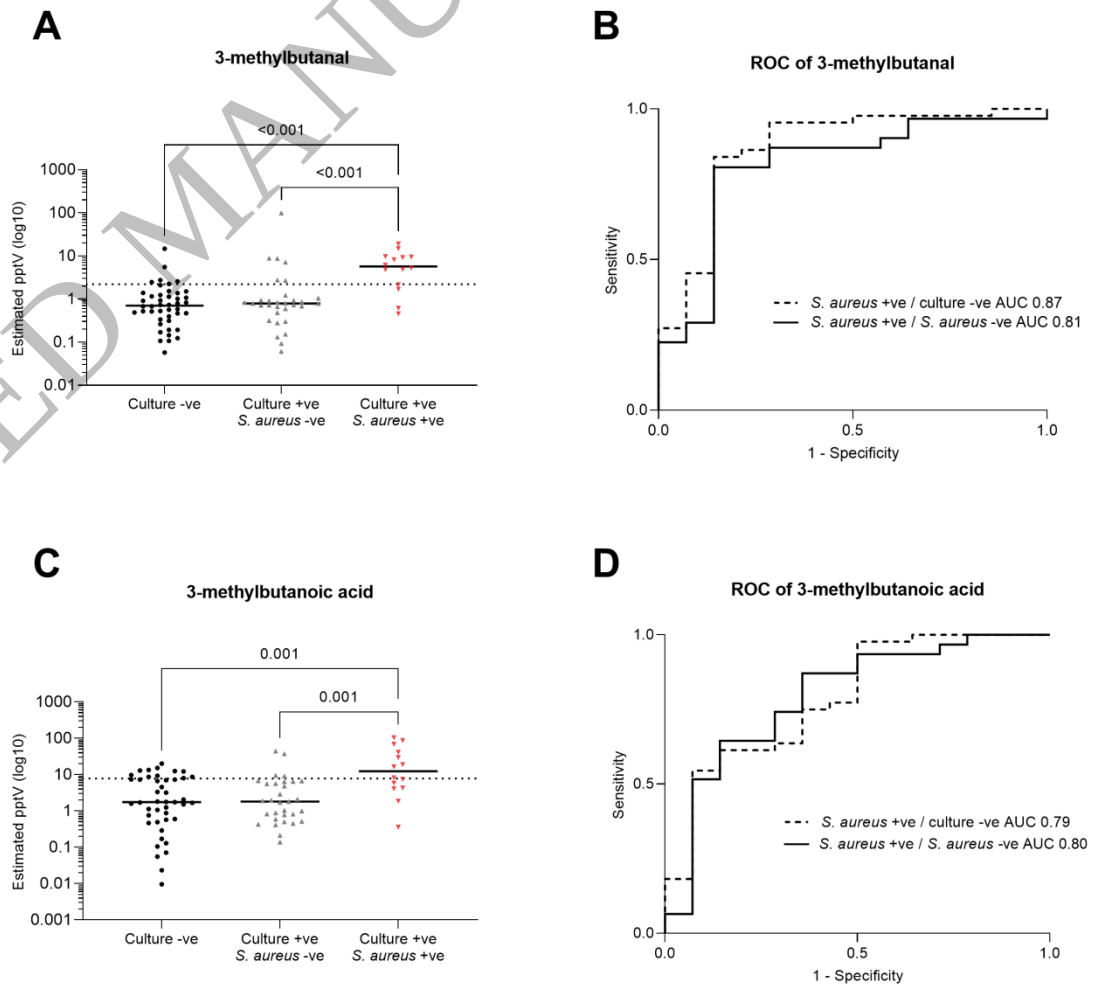

Figure 2
339x190 mm (x DPI)