Octane in exhaled breath to diagnose acute respiratory distress syndrome in invasively ventilated intensive care unit patients

Laura A. Hagens [1], Nanon F.L. Heijnen [2], Marry R. Smit [1], Alwin R.M. Verschueren [3], Tamara M.E. Nijsen [3], Inge Geven [3], Cristian N. Presură [3], Ronald Rietman [3], Dominic W. Fenn [4,5], Paul Brinkman [4], Marcus J. Schultz [1,6,7], Dennis C.J.J. Bergmans [2,8], Ronny M. Schnabel [2] and Lieuwe D.J. Bos [1,4,5] on behalf of the DARTS consortium

[1]Department of Intensive Care, Amsterdam UMC location University of Amsterdam, Amsterdam, The Netherlands. [2]Department of Intensive Care, Maastricht University Medical Centre+, Maastricht, The Netherlands. [3]Sleep and Respiratory Solutions, Philips Research, Eindhoven, The Netherlands. [4]Department of Respiratory Medicine, Amsterdam UMC location University of Amsterdam, Amsterdam, The Netherlands. [5]Laboratory of Experimental Intensive Care and Anesthesiology (LEICA), Amsterdam UMC location University of Amsterdam, Amsterdam, The Netherlands. [6]Mahidol-Oxford Tropical Medicine Research Unit (MORU), Mahidol University, Bangkok, Thailand. [7]Nuffield Department of Medicine, University of Oxford, Oxford, UK. [8]School of Nutrition and Translational Research in Metabolism (NUTRIM), Maastricht University, Maastricht, The Netherlands.

Corresponding author: Laura A. Hagens (l.a.hagens@amsterdamumc.nl)

Shareable abstract (@ERSpublications)
Octane in exhaled breath has poor diagnostic accuracy for ARDS. This disqualifies the use of octane as a biomarker in the diagnosis of ARDS and challenges most of the research performed to date in the field of exhaled breath metabolomics. https://bit.ly/3OSf8iC

Cite this article as: Hagens LA, Heijnen NFL, Smit MR, *et al*. Octane in exhaled breath to diagnose acute respiratory distress syndrome in invasively ventilated intensive care unit patients. *ERJ Open Res* 2023; 9: 00214-2023 [DOI: 10.1183/23120541.00214-2023].

Received: 5 April 2023
Accepted: 3 Aug 2023

Abstract

Background The concentration of exhaled octane has been postulated as a reliable biomarker for acute respiratory distress syndrome (ARDS) using metabolomics analysis with gas chromatography and mass spectrometry (GC-MS). A point-of-care (POC) breath test was developed in recent years to accurately measure octane at the bedside. The aim of the present study was to validate the diagnostic accuracy of exhaled octane for ARDS using a POC breath test in invasively ventilated intensive care unit (ICU) patients.

Methods This was an observational cohort study of consecutive patients receiving invasive ventilation for at least 24 h, recruited in two university ICUs. GC-MS and POC breath tests were used to quantify the exhaled octane concentration. ARDS was assessed by three experts following the Berlin definition and used as the reference standard. The area under the receiver operating characteristic curve (AUC) was used to assess diagnostic accuracy.

Results 519 patients were included and 190 (37%) fulfilled the criteria for ARDS. The median (interquartile range) concentration of octane using the POC breath test was not significantly different between patients with ARDS (0.14 (0.05–0.37) ppb) and without ARDS (0.11 (0.06–0.26) ppb; p=0.64). The AUC for ARDS based on the octane concentration in exhaled breath using the POC breath test was 0.52 (95% CI 0.46–0.57). Analysis of exhaled octane with GC-MS showed similar results.

Conclusions Octane in exhaled breath has insufficient diagnostic accuracy for ARDS. This disqualifies the use of octane as a biomarker in the diagnosis of ARDS and challenges most of the research performed up to now in the field of exhaled breath metabolomics.

Introduction

Acute respiratory failure requiring invasive mechanical ventilation is frequently caused by acute respiratory distress syndrome (ARDS), which is characterised by protein-rich pulmonary oedema [1]. ARDS affects ~10% of patients admitted to the intensive care unit (ICU) and is associated with a mortality rate of 40% [2]. The presence of ARDS is established based on the criteria formulated in the Berlin definition [3], which includes timing, gas exchange abnormalities and bilateral infiltrates on chest radiography indicative of pulmonary oedema, not completely explained by a cardiogenic cause [3]. Despite these criteria, the diagnosis

is frequently missed, leading to inadequate treatment and ventilation strategies in these patients [2, 4]. Current tests to diagnose ARDS are neither sensitive nor specific [5] and there remains a need for more objective bedside identification of these patients, *e.g.* using biomarkers [6].

Biological markers can be sampled from various biospecimens, but availability of pulmonary samples is limited due to sampling issues. Breath analysis could provide a solution, as it contains hundreds of different volatile organic compounds (VOCs) that are reflective of metabolic changes in the lungs [7]. Breath molecules can be separated, quantified and identified using gas chromatography and mass spectrometry (GC-MS). This technique was previously used to discover and validate a three-metabolite diagnostic model with good accuracy for ARDS [8], with breath octane being the most important biomarker explaining the full diagnostic accuracy. Endogenous octane formation results from peroxidation of oleic acid [9], and patients with ARDS have an increased concentration of oleic acid and higher levels of oxidative stress, providing a strong biochemical rationale for the link between octane and ARDS [10, 11]; a link that remains uncertain for the other two metabolites [8].

Exhaled breath analysis by GC-MS is not feasible to implement in clinical practice. GC-MS requires specialised personnel and equipment, making this technique unsuitable to be used as a rapid bedside test. To overcome the disadvantages of GC-MS, our team recently developed a point-of-care (POC) breath test for octane, which is highly accurate at ppb concentration [12]. The POC breath test can selectively quantify octane concentrations in exhaled breath and provides a result within hours [12]. Besides the technological advances, the POC breath test still required further validation in prospective studies, as a step to implementation into clinical practice.

The aim of this study was to validate the concentration of octane in exhaled breath as a reliable biomarker for ARDS. For this purpose, we tested both the novel POC breath test and GC-MS as index tests in invasively ventilated ICU patients. Additionally, breath octane concentrations were also compared with radiological measures of pulmonary oedema.

Methods

This was a multicentre observational cohort study with longitudinal sampling of consecutive patients admitted to the ICU receiving invasive ventilation for at least 24 h. The protocol of this study was previously published [13]. Patients were recruited from 26 March 2019 until 1 March 2021, in two academic hospitals in the Netherlands (Amsterdam UMC location University of Amsterdam, Amsterdam and Maastricht University Medical Centre+, Maastricht). The Institutional Review Boards of both centres waived the need for ethical approval of the protocol (W18_311#18.358 and 2019-1137). The trial was registered at the Netherlands Trial Register (NL8226) with the tag “DARTS study”. Details on the informed consent procedure were described previously [13].

Study population

Inclusion criteria were an expected ventilation duration of at least 24 h and age ⩾18 years. Patients were excluded when they received invasive ventilation for >48 h at any moment in the 7 days preceding the moment of inclusion or when they were expected to decease within 24 h. Tracheotomised patients were also excluded. Finally, patients were excluded if breath sampling was deemed clinically inappropriate (*e.g.* in patients with highly contagious conditions) or if consent was withdrawn by the patient or family.

Study procedures

Patients were recruited within the first 48 h after the start of invasive ventilation (supplementary figure S1) and two assessments were performed 24 h apart (time-points 1 and 2). Both assessments consisted of clinical data collection, breath sampling and lung ultrasound (LUS) examination (supplementary figure S1). The chest radiograph and worst arterial oxygen tension (P_{aO_2})/inspiratory oxygen fraction (F_{IO_2}) ratio in the 24 h before breath sampling were collected and used for ARDS diagnosis. If available, a chest computed tomography (CT) scan was collected within 72 h of inclusion. In patients suspected of ARDS during time-points 1 and 2, additional breath samples were taken at approximately 48 h, 96 h and 12 days after inclusion (supplementary figure S1).

ARDS diagnosis: reference test

A panel of three experts, clinicians with expertise in the field of diagnosing ARDS, independently reviewed the available clinical variables and imaging to determine the ARDS diagnosis. This procedure of diagnosing ARDS was described previously [14]. In short, each expert scored the chest radiograph and chest CT scan on 1) the presence of bilateral opacities consistent with ARDS, 2) the most likely aetiology of the opacities and 3) if cardiogenic oedema was the most likely explanation for the abnormalities [3].

Thereafter the expert decided, based on all available information, whether the patient fulfilled the criteria of the Berlin definition for ARDS. The confidence of this diagnosis was scored on an 8-grade scale (supplementary figure S2a), leading to a grade ranging from 1 to 8, respectively corresponding to high confidence of "no ARDS" or high confidence of "ARDS" on both extremes [15].

Using an algorithm (supplementary figure S2b) that combined the confidence grades of the experts, for each imaging modality and day, the diagnosis was divided into three categories: 1) "no ARDS" if sufficient confidence was reached to exclude ARDS, 2) "ARDS" if sufficient confidence was reached to diagnose ARDS and 3) "uncertain diagnosis" in case of general uncertainty (none of the experts had strong confidence) or conflicting results. To reach a final diagnosis, we prioritised imaging with the most accurate information available for classification (supplementary figure S2c). For the "uncertain diagnosis" category, a consensus meeting was held to determine the final classification for the patient, resulting in two additional categories: "likely ARDS" and "likely no ARDS". Additionally, the amount of pulmonary oedema was estimated using the Radiographic Assessment of Lung Oedema (RALE) and global LUS scores, as described in the supplementary material. Assessment of the reference test was blinded for any results from the index test.

Breath sampling

Exhaled breath was drawn for 6 min using a dedicated breath gas sampler and polytetrafluoroethylene tubing (Swagelok, Warrington, UK) [12], by connecting a three-way stopcock to the expiratory limb in the case of double-limb ventilation and after the heat–moisture exchange filter on the ventilator site in the case of single-limb ventilation (supplementary figure S3). This method has been shown to be reliable and reproducible [12, 16]. VOCs in exhaled breath were absorbed onto two sorbent tubes filled with 300 mg Carbograph 5TD (Markes International, Llantrisant, UK) and 90 mg Tenax GR (Sigma-Aldrich Chemie, Zwijndrecht, The Netherlands). Subsequently the sorbent tubes were refrigerated until analysis.

POC breath test: index test

The POC breath test was specifically designed and validated for the present study aiming to quantify the amount of (n-)octane present in exhaled breath samples of ICU patients [12]. The POC breath test consists of a thermal desorption unit, a GC column and a photo-ionisation detector. The retention time of octane and sensitivity of the detector were recalibrated at regular intervals using a standardised gas mixture. This was used to calculate the concentration of octane (ppb). The reader of the octane concentration was blinded to the ARDS diagnosis.

GC-MS: additional index test

To validate the amount of octane measured with the POC breath test, the other sorbent tube was analysed by means of GC-MS as described previously [17]. GC-MS analysis, de-noising, peak detection and alignment were performed using the R xcms package (Scripps Center for Metabolomics, La Jolla, CA, USA). The concentration of octane (ppb) was calculated with the *m/z* 85 peak using a calibration curve generated with a standardised gas mixture. Again, the reader was blinded to the ARDS diagnosis.

Study outcomes

The primary end-point of this study was the diagnostic accuracy of octane concentration in exhaled breath for ARDS, using the ARDS diagnosis by the panel of experts as the reference test. Secondary end-points were: 1) the added diagnostic value of the POC breath test on top of the Lung Injury Prediction Score (LIPS), 2) the association between the octane concentration and the certainty of the ARDS diagnosis (likely or certain diagnoses), and 3) changes in concentration of exhaled octane in patients with ARDS.

As a post-hoc analysis the correlations of the octane concentration with radiological findings of pulmonary oedema were assessed. This was performed for the global LUS score and the RALE score.

Sample size justification

With an expected sensitivity of 80% [8] and a minimally acceptable lower confidence limit of 65%, at least 52 patients with ARDS were required for the study [18]. At a predicted incidence of 10.4% [2], at least 500 patients were needed for the primary end-point.

Statistical analysis

All statistical analyses were performed in R version 4.0.3 (www.r-project.org) using the RStudio interface. For the primary outcome the area under the receiver operating characteristic curve (AUC) was calculated with the patients divided into ARDS ("likely ARDS" and "ARDS") or no ARDS ("likely no ARDS" and "no ARDS"). The maximum octane concentration of the first two time-points was used. Data were summarised based on the type of data and the distribution of the values for continuous variables.

Categorical variables were presented as number and percentage, and differences between the ARDS and no ARDS group and the ARDS categories were tested with a Chi-squared test. Continuous variables were presented as mean with standard deviation for variables with a normal distribution (based on histograms, before and after transformation) or with a presumed normal distribution (such as height) and differences between groups were tested with a t-test or one-way ANOVA, as appropriate. Continuous variables without a normal distribution were presented as median with interquartile range (IQR) and differences between groups were tested with a Mann–Whitney U-test or Kruskal–Wallis test, as appropriate. Receiver operating characteristic curve analysis was performed using the pROC package [19]. Correlation between two continuous variables was performed with the Pearson correlation for normally distributed variables and with the Spearman correlation for non-normally distributed variables. A p-value <0.05 was considered statistically significant.

Results

Patients

A total of 519 patients were included into the analysis (figure 1). 37% of the included patients (190 out of 519) fulfilled the criteria for ARDS. In a total of 158 out of 519 patients (30%) a consensus meeting was required to reach a diagnosis. Consensus classified 77 patients (49%) as having ARDS (category "likely ARDS") and 81 (51%) as not having ARDS (category "likely no ARDS") (supplementary table S2). Patient characteristics are presented in table 1 and reflect typical differences, such as a lower P_{aO_2}/F_{IO_2} and compliance of the respiratory system, higher positive end-expiratory pressure and higher pulmonary oedema scores in patients diagnosed with ARDS.

Technical failure led to 27 missing POC breath test measurements and 29 missing GC-MS measurements at time-point 1. At time-point 2, 320 POC breath test measurements and 303 GC-MS measurements were available. GC-MS measurements were missing on both time-points for 11 patients in the ARDS group and for nine patients in the group without ARDS. POC breath test results were missing in 10 patients, all in the group without ARDS.

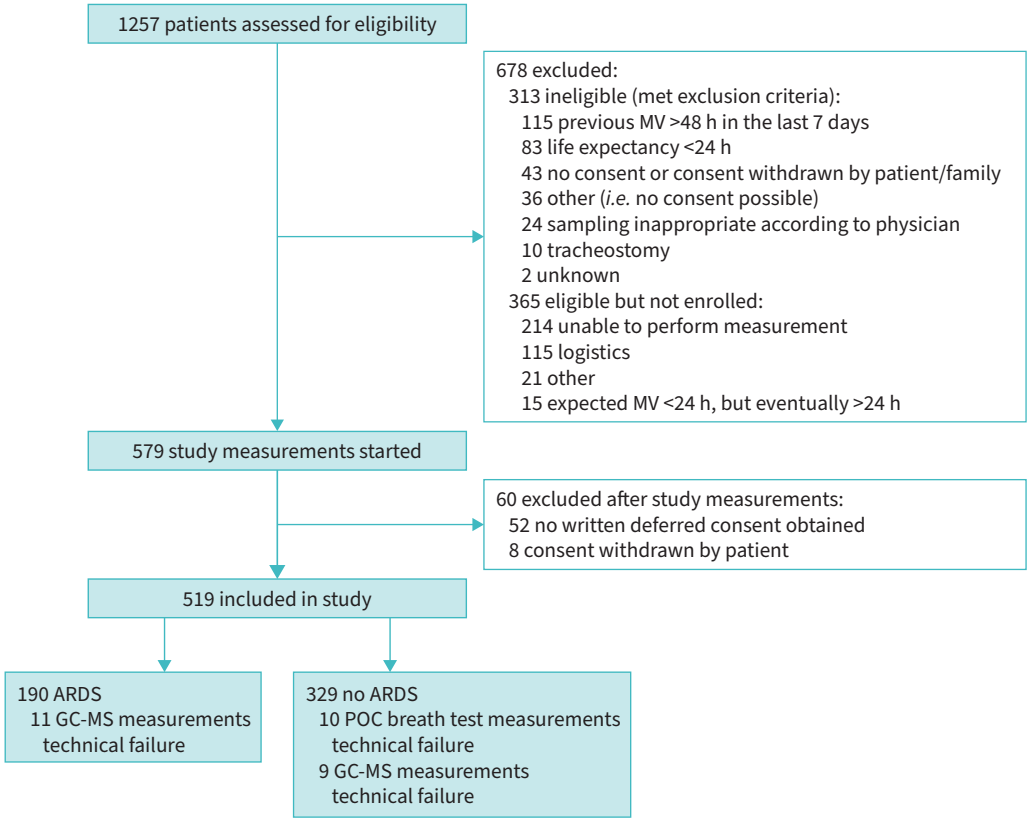

FIGURE 1 Study flowchart. MV: mechanical ventilation; ARDS: acute respiratory distress syndrome; GC-MS: gas chromatography-mass spectrometry; POC: point-of-care.

TABLE 1 Patient characteristics

	No ARDS (n=329)	ARDS (n=190)	p-value
Age (years)	62±15.4	63±12.9	0.43
Male	223 (68)	131 (69)	0.86
BMI ($kg \cdot m^{-2}$)	26.1 (23.0–30.0)	27.0 (24.1–30.3)	0.15
Admission characteristics			
Admission type			<0.001
Emergency surgical	60 (18.2)	15 (7.9)	
Medical	220 (66.9)	161 (84.7)	
Planned surgical	49 (14.9)	14 (7.4)	
ARDS cause			
Non-pulmonary		48 (25.3)	
Pulmonary		142 (74.7)	
ARDS severity			
Mild		26 (13.7)	
Moderate		98 (51.6)	
Severe		65 (34.2)	
Pneumonia	37 (11.2)	126 (66.3)	<0.001
APACHE II score	21 (15–26)	20 (15–24)	0.012
SOFA score	9 (8–11)	9 (7–11)	0.033
LIPS	4.5 (3.0–6.5)	6.0 (4.5–7.5)	<0.001
Global LUS score	5 (2–9)	13 (8–16)	<0.001
RALE score	12 (7–18)	20 (15–28)	<0.001
COVID-19	2 (0.6)	63 (33.2)	<0.001
Duration of hospital stay before ICU (days)	1.0 (1.0–3.0)	2.0 (1.0–5.8)	<0.001
Ventilation and gas exchange			
MV duration at day 1 (h)	21 (14–31)	21 (12–29)	0.66
Compliance ($mL \cdot cmH_2O^{-1}$)	34.9 (25.5–49.4)	28.8 (21.6–42.3)	0.001
P_{aO_2}/F_{IO_2} [#] (mmHg)	233 (154–319)	118 (86–165)	<0.001
PEEP (cmH_2O)	6 (5–8)	10 (8–12)	<0.001
V_T PBW ($mL \cdot kg^{-1}$)	7.07 (6.16–8.53)	7.24 (5.92–8.49)	0.81
V_T (mL)	464 (398–540)	458 (384–559)	0.86
Driving pressure (cmH_2O)	13 (9–16)	15 (10–19)	0.001
Respiratory rate (min^{-1})	18 (15–22)	20 (16–25)	<0.001
P_{max} (cmH_2O)	19 (15–24)	25 (19–29)	<0.001
Outcomes			
Hospital LOS (days)	17 (8–31)	20 (12–31)	0.023
ICU LOS (days)	6 (3–11)	9 (4–16)	<0.001
ICU mortality	101 (30.7)	71 (37.4)	0.26
30-day mortality	121 (36.8)	78 (41.1)	0.40

Data presented as mean±SD, n (%) or median (interquartile range), unless otherwise stated. ARDS: acute respiratory distress syndrome; BMI: body mass index; APACHE II: Acute Physiology and Chronic Health Evaluation II; SOFA: Sequential Organ Failure Assessment; LIPS: Lung Injury Prediction Score; LUS: lung ultrasound; RALE: Radiographic Assessment of Lung Oedema; COVID-19: coronavirus disease 2019; ICU: intensive care unit; MV: mechanical ventilation; P_{aO_2}: arterial oxygen tension; F_{IO_2}: inspiratory oxygen fraction; PEEP: positive end-expiratory pressure; V_T: tidal volume; PBW: predicted body weight; P_{max}: maximal pressure; LOS: length of stay. [#]: P_{aO_2}/F_{IO_2} is defined as the lowest P_{aO_2}/F_{IO_2} in the 24 h before day 1. p-values were calculated using the Chi-squared, t-test or Mann–Whitney U-test depending on the type and distribution of the variable.

Octane concentration in exhaled breath

In the ARDS group, the median (IQR) concentration of octane was 0.14 (0.05–0.37) ppb using the POC breath test and 0.14 (0.05–0.37) ppb using GC-MS; in the group without ARDS, the median (IQR) concentration of octane was 0.11 (0.06–0.26) ppb using the POC breath test and 0.15 (0.08–0.34) ppb using GC-MS (figure 2). There was no significant difference in breath octane concentration between patients with and without ARDS measured with the POC breath test (p=0.64) or GC-MS (p=0.75). There was no significant difference in exhaled octane concentration between any of the ARDS categories related to the certainty of diagnosis (certain ARDS *versus* certain no ARDS, p=0.40 and p=0.94, respectively) (supplementary figure S4). All breath octane concentrations can be found in the supplementary material, together with values for specific subgroups (supplementary tables S3, S5 and S6).

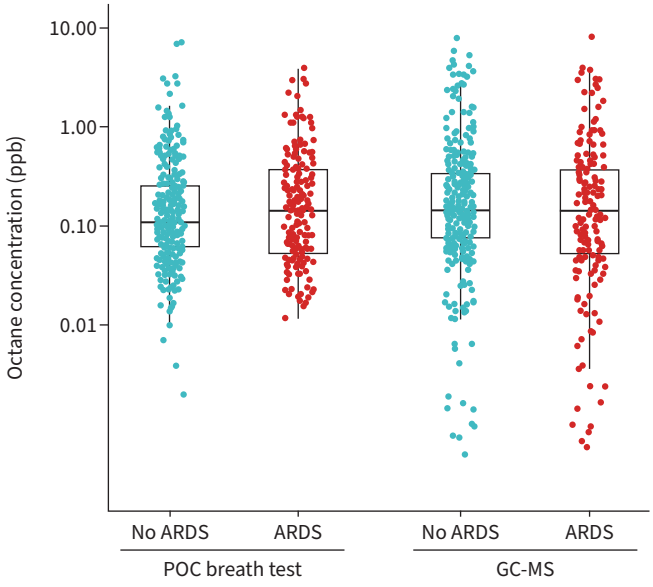

FIGURE 2 Exhaled breath octane concentration (log_{10} scale) compared between patients with and without acute respiratory distress syndrome (ARDS) measured with the point-of-care (POC) breath test and gas chromatography-mass spectrometry (GC-MS) displayed stratified for ARDS (certain "ARDS" and "likely ARDS" *versus* certain "no ARDS" and "likely no ARDS", dichotomised). There was no significant difference in breath octane concentration between patients with and without ARDS for the POC breath test (p=0.64) or GC-MS (p=0.75).

Diagnostic accuracy

The AUC for ARDS based on the octane concentration in exhaled breath using the POC breath test was 0.52 (95% CI 0.46–0.57), which corresponds to a bad diagnostic accuracy (figure 3a) [20]. The AUC using GC-MS showed similar results (0.53, 95% CI 0.48–0.59) (figure 3b). LIPS, a clinical prediction

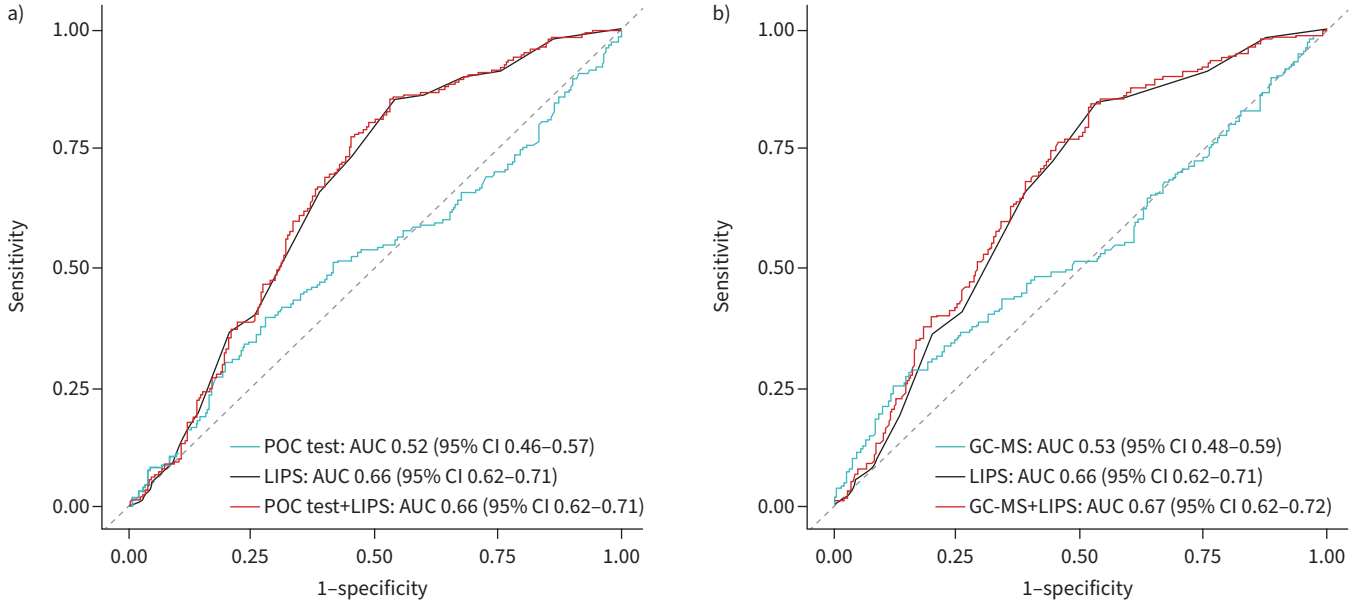

FIGURE 3 Area under the receiver operating characteristic curve (AUC) describing the diagnostic accuracy for acute respiratory distress syndrome of the octane concentration measured with a) point-of-care (POC) breath test and Lung Injury Prediction Score (LIPS) and b) gas chromatography-mass spectrometry (GC-MS) and LIPS.

score, had a better AUC for ARDS (0.66, 95% CI 0.62–0.71), corresponding to moderate diagnostic accuracy. Combining LIPS with the octane concentration measured with the POC breath test resulted in an AUC of 0.66 (95% CI 0.62–0.71) and combining LIPS with the octane concentration derived from GC-MS resulted in an AUC of 0.67 (95% CI 0.62–0.72).

Longitudinal sampling

To understand if there were longitudinal differences in breath octane concentration between patients with and without ARDS, we separated time-points 1 and 2 in a secondary analysis (figure 4) and found no differences on either time-point (time-point 1: p=0.63; time-point 2: p=0.99). When extending the observation period to further time-points in patients diagnosed with ARDS, we did not observe any trends over time (p-values compared with time-point 1: time-point 2: p=0.30; time-point 3: p=0.55; time-point 4: p=0.10; time-point 5: p=0.07) (figure 4 and supplementary table S4).

Correlation with radiological measures of pulmonary oedema

The octane concentration in exhaled breath had no correlation with the global LUS score (ρ: 0.03, 95% CI −0.06–0.13) and a weak correlation with the RALE score (ρ: 0.11, 95% CI 0.01–0.20).

Discussion

Despite the strong pre-clinical rationale and a promising discovery study, the diagnostic accuracy of the octane concentration in exhaled breath of invasively ventilated ICU patients measured with a POC breath test was bad for the early diagnosis of ARDS. This was not due to measurement error, as similar results were found when the gold standard GC-MS test was used and the POC breath test has previously been shown to reliably detect octane [12]. The clinical certainty of ARDS diagnosis defined by a panel of experts also did not drive the poor accuracy as breath octane concentrations were similar across the certainties of diagnosis. There also was a poor correlation between the measured octane concentration in exhaled breath and pulmonary oedema scores based on chest radiography and LUS. Together, the results of this study disqualify octane in exhaled breath as a diagnostic test for ARDS.

We were not able to replicate the findings of an earlier study performed by our group, which revealed octane as a good diagnostic marker of ARDS [8]. The current study is more than five times larger than any

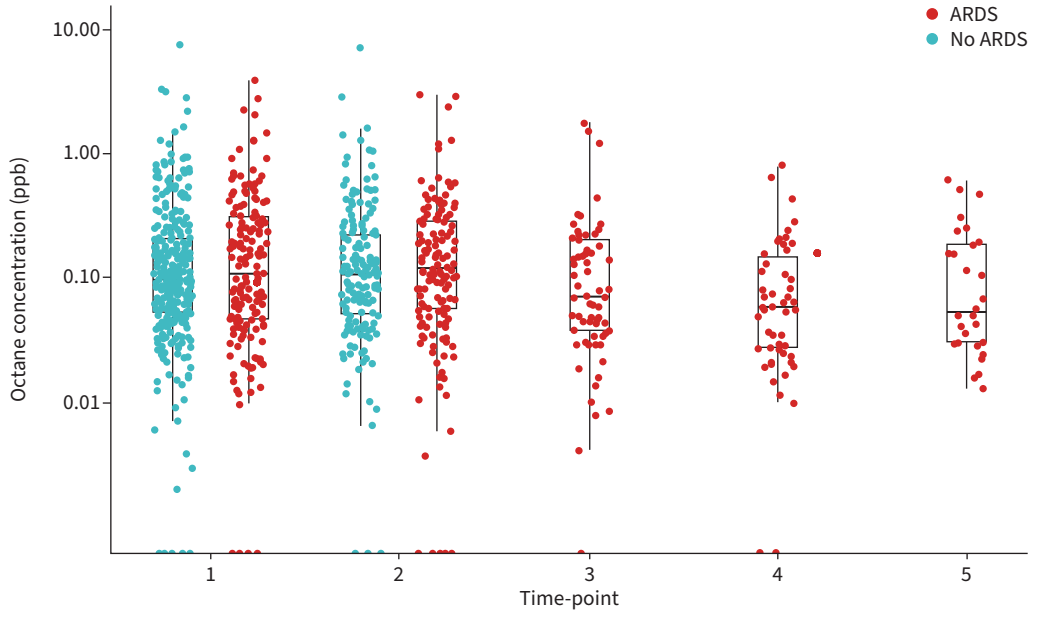

FIGURE 4 Longitudinal sampling. Octane concentrations (log_{10} scale) as measured with the point-of-care breath test. Time-points 1 and 2: both the acute respiratory distress syndrome (ARDS) and no ARDS groups. Time-points 3–5: only ARDS patients. Time-point 3 corresponds to day 3 after inclusion, time-point 4 to day 5 and time-point 5 to day 12. p-values compared with time-point 1: time-point 2: p=0.30; time-point 3: p=0.55; time-point 4: p=0.10; time-point 5: p=0.07.

previous study in the field of breath research for ARDS diagnosis [8, 21, 22]. This results in a more valid estimation of the diagnostic accuracy. Data were collected from consecutive patients, limiting selection bias and thus providing an unbiased diagnostic accuracy estimate. As in previous studies, the timing of the sampling was early in the disease course. Secondary analyses focused on time-dependent changes that did not reveal improvements in accuracy. The use of an expert panel to establish the reference standard further limits bias in the setting of an imperfect reference standard, another improvement over previous studies [23]. Taken together, the study presented here has methodological advantages over previous studies and provides a definitive answer against breath octane as a diagnostic marker for ARDS.

So why did previous studies observe a difference in exhaled breath metabolites between patients with and without ARDS? The most compelling explanation is that these were false discoveries. Exhaled breath metabolomics, like most "omics" technologies, is a statistical challenge because one performs untargeted discovery on a larger number of predictors than the number of clinical cases. Our group has previously shown that exhaled breath research is particularly prone to overfitting results to the training set [24]. However, octane was validated in a temporal validation cohort and did perform well under those circumstances. We can exclude the influence of sampling technique, which was the same in both studies, measurement technique, which was highly controlled and validated as well, and diagnostic criteria of the reference standard. Therefore, we propose that the diagnostic accuracy of octane in the previous study should be considered a chance finding despite external validation.

This study has particular strengths, as outlined, but also some limitations. Around 4% of the measurements failed and this may have had a small influence on the results. However, this percentage of missingness is insufficient to bias the observed diagnostic accuracy in such a way that it would alter our conclusions. There is also no gold standard diagnosis for ARDS and we have to acknowledge that this leaves room for interpretation of the Berlin definition. We tried to limit the influence of individual observers and therefore bias caused by poor inter-rater agreement on the reference standard by using an expert panel, but we cannot exclude the possibility that some patients would have been classified differently by other clinicians. Considering that there were no differences in exhaled breath octane concentration between patients who certainly had no ARDS and certainly had ARDS, excluding all cases that required a consensus meeting to come to a definitive diagnosis, we conclude that this would also not influence our result in any meaningful way. We did not evaluate other breath biomarkers besides octane and future explorations of this dataset may provide novel ARDS breath signatures. Assessing a panel of biomarkers might be more beneficial compared with a single one [6]. Finally, ARDS is a heterogeneous condition and exhaled breath metabolomics might still facilitate the identification of subphenotypes rather than diagnosis.

The results of this study are a cautionary tale to the breath metabolomics field, with much relevance to all "omics". We describe a decade-long journey from discovery to validation resulting in the disqualification of a breath biomarker with a good biochemical rationale and very promising discovery data based on a large, unbiased validation. Future studies are needed to assess the diagnostic accuracy of a combination of breath biomarkers for the diagnosis of ARDS, as the analyses presented here cannot provide data on this topic. The POC breath test was shown to measure octane reliably during the study period, which is a major step towards feasibility of this POC breath test. Technologies like this are an important enabler for larger scale breath analysis studies in the future that are needed. For many areas of research, such thorough validation is not sought and our results suggest that these findings should be treated with the utmost caution.

Conclusions

We attempted to validate exhaled breath octane as diagnostic marker for ARDS, but revealed it has insufficient diagnostic accuracy. This disqualifies this biomarker from being used in the diagnosis of ARDS and challenges most of the research performed up to now in the field of exhaled breath metabolomics.

Provenance: Submitted article, peer reviewed.

DARTS consortium: Laura A. Hagens, Marry R. Smit, Marcus J. Schultz and Lieuwe D.J. Bos (Amsterdam UMC); Nanon F.L. Heijnen, Dennis C.J.J. Bergmans and Ronny M. Schnabel (Maastricht UMC); and Alwin R.M. Verschueren, Tamara M.E. Nijsen, Inge Geven and Cristian N. Presură (Philips Research).

This study is registered at www.onderzoekmetmensen.nl with identifier number NL8226.

Conflict of interest: L.D.J. Bos reports grants from the Dutch Lung Foundation (Young Investigator grant and Dirkje Postma Award), the Dutch Lung Foundation and Health Holland (public–private partnership grant), and the IMI

COVID19 initiative, and an Amsterdam UMC fellowship, a ZonMW COVID-19 Urgency grant and the ERS Gold Metal for ARDS; he reports participating in advisory boards for Sobi, Exvastat, Santhera, Pfizer and AstraZeneca, all paid to his institution, and consultancy for Scailyte, Santhera and Janssen & Janssen, all paid to his institution, outside the submitted work. A.R.M. Verschueren, T.M.E. Nijsen, I. Geven, C.N. Presură and R. Rietman are employees of Philips Research. L.A. Hagens, N.F.L. Heijnen, M.R. Smit, D.W. Fenn, P. Brinkman, M.J. Schultz, D.C.J.J. Bergmans and R.M. Schnabel have no conflicts of interest to declare.

Support statement: This study was funded by Health Holland and the Dutch Lung Fund. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Mac Sweeney R, McAuley DF. Acute respiratory distress syndrome. *Lancet* 2016; 388: 2416–2430.
2 Bellani G, Laffey JG, Pham T, *et al.* Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016; 315: 788–800.
3 ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012; 307: 2526–2533.
4 Bellani G, Pham T, Laffey JG. Missed or delayed diagnosis of ARDS: a common and serious problem. *Intensive Care Med* 2020; 46: 1180–1183.
5 Binnie A, Tsang JLY, dos Santos CC. Biomarkers in acute respiratory distress syndrome. *Curr Opin Crit Care* 2014; 20: 47–55.
6 Hagens LA, Heijnen NFL, Smit MR, *et al.* Systematic review of diagnostic methods for acute respiratory distress syndrome. *ERJ Open Res* 2020; 7: 00504-2020.
7 Moser B, Bodrogi F, Eibl G, *et al.* Mass spectrometric profile of exhaled breath – field study by PTR-MS. *Respir Physiol Neurobiol* 2005; 145: 295–300.
8 Bos LDJ, Weda H, Wang Y, *et al.* Exhaled breath metabolomics as a noninvasive diagnostic tool for acute respiratory distress syndrome. *Eur Respir J* 2014; 44: 188–197.
9 Kneepkens CMF, Lepage G, Roy CC. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. *Free Radic Biol Med* 1994; 17: 127–160.
10 Quinlan GJ, Lamb NJ, Evans TW, *et al.* Plasma fatty acid changes and increased lipid peroxidation in patients with adult respiratory distress syndrome. *Crit Care Med* 1996; 24: 241–246.
11 Carpenter CT, Price PV, Christman BW. Exhaled breath condensate isoprostanes are elevated in patients with acute lung injury or ARDS. *Chest* 1998; 114: 1653–1659.
12 Hagens LA, Verschueren ARM, Lammers A, *et al.* Development and validation of a point-of-care breath test for octane detection. *Analyst* 2021; 146: 4605–4614.
13 Hagens LA, Heijnen NFL, Smit MR, *et al.* Diagnosis of acute respiratory distress syndrome (DARTS) by bedside exhaled breath octane measurements in invasively ventilated patients: protocol of a multicentre observational cohort study. *Ann Transl Med* 2021; 9: 1262.
14 Hagens LA, Van der Ven FLIM, Heijnen NFL, *et al.* Improvement of an interobserver agreement of ARDS diagnosis by adding additional imaging and a confidence scale. *Front Med* 2022; 9: 950827.
15 Sjoding MW, Hofer TP, Co I, *et al.* Interobserver reliability of the Berlin ARDS definition and strategies to improve the reliability of ARDS diagnosis. *Chest* 2018; 153: 339–348.
16 Bos LDJ, Wang Y, Weda H, *et al.* A simple breath sampling method in intubated and mechanically ventilated critically ill patients. *Respir Physiol Neurobiol* 2014; 191: 67–74.
17 Bos LDJ, Van Walree IC, Kolk AHJ, *et al.* Alterations in exhaled breath metabolite-mixtures in two rat models of lipopolysaccharide-induced lung injury. *J Appl Physiol* 2013; 115: 1487–1495.
18 Flahault A, Cadilhac M, Thomas G. Sample size calculation should be performed for design accuracy in diagnostic test studies. *J Clin Epidemiol* 2005; 58: 859–862.
19 Robin X, Turck N, Hainard A, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; 12: 77.
20 Šimundić A-M. Measures of diagnostic accuracy: basic definitions. *EJIFCC* 2009; 19: 203–211.
21 Zhou M, Sharma R, Zhu H, *et al.* Rapid breath analysis for acute respiratory distress syndrome diagnostics using a portable two-dimensional gas chromatography device. *Anal Bioanal Chem* 2019; 411: 6435–6447.
22 Schubert JK, Müller WP, Benzing A, *et al.* Application of a new method for analysis of exhaled gas in critically ill patients. *Intensive Care Med* 1998; 24: 415–421.
23 Walsh T. Fuzzy gold standards: approaches to handling an imperfect reference standard. *J Dent* 2018; 74: Suppl. 1, S47–S49.
24 Leopold JH, Bos LDJ, Sterk PJ, *et al.* Comparison of classification methods in breath analysis by electronic nose. *J Breath Res* 2015; 9: 046002.