ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original article

Breath analysis for detection and trajectory monitoring of acute respiratory distress syndrome in swine

Ruchi Sharma, Menglian Zhou, Mohamad Hakam Tiba, Brendan M. McCracken, Robert P. Dickson, Christopher E. Gillies, Michael W. Sjoding, Jean A. Nemzek, Kevin R. Ward, Kathleen A. Stringer, Xudong Fan

Please cite this article as: Sharma R, Zhou M, Tiba MH, *et al*. Breath analysis for detection and trajectory monitoring of acute respiratory distress syndrome in swine. *ERJ Open Res* 2021; in press (https://doi.org/10.1183/23120541.00154-2021).

Breath Analysis for Detection and Trajectory Monitoring of Acute Respiratory Distress Syndrome in Swine

Ruchi Sharma[1], Menglian Zhou[1], Mohamad Hakam Tiba[2,6], Brendan M. McCracken[2,6], Robert P. Dickson[3,6], Christopher E. Gillies [2,6,8], Michael W. Sjoding[3,6], Jean A. Nemzek[4,5,6], Kevin R. Ward[1,2,6], Kathleen A. Stringer[3,6,7], and Xudong Fan[1,6,+]

[1]Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, U.S.A.

[2]Department of Emergency Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[3]Department of Internal Medicine, Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[4]Unit of Laboratory Animal Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[5]Department of Pathology, University of Michigan, Ann Arbor, MI, U.S.A.

[6]Michigan Center for Integrative Research in Critical Care, University of Michigan, Ann Arbor, MI, U.S.A.

[7]Department of Clinical Pharmacy, College of Pharmacy, University of Michigan, Ann Arbor, MI, U.S.A

[8]Michigan Institute for Data Science (MIDAS), Office of Research, University of Michigan, Ann Arbor, MI, U.S.A

[+]xsfan@umich.edu

Abstract

Despite the enormous impact on human health, acute respiratory distress syndrome (ARDS) is ill-defined, and its timely diagnosis is difficult, as is tracking the course of the syndrome. The objective of this pilot study was to explore the utility of breath collection and analysis methodologies to detect ARDS through changes in the volatile organic compound (VOC) profiles present in breath. Five male Yorkshire mix swine were studied and ARDS was induced utilizing both direct and indirect lung injury. An automated portable gas chromatography device developed in-house was used for point of care breath analysis and to monitor swine breath hourly, starting from the initiation of the experiment until the development of ARDS, which was adjudicated based on the Berlin criteria at the breath sampling points and confirmed by lung biopsy at the end of the experiment. A total of 67 breath samples (chromatograms) were collected and analyzed. Through machine learning, principal component analysis, and linear discrimination analysis, seven VOCs biomarkers were identified that distinguished ARDS. These represent seven of the nine biomarkers found in our breath analysis study of human ARDS corroborating our findings. We also demonstrated that breath analysis detects changes 1-6 hours earlier than the clinical adjudication based on the Berlin criteria. The findings provide proof of concept that breath analysis can be used for the identification of early changes associated with ARDS pathogenesis in swine. Its clinical application could provide intensive care clinicians with a non-invasive diagnostic tool for early detection and continuous monitoring of ARDS.

Keywords: Lung injury, Volatile organic compound, Gas chromatography, Large animal model, Machine learning

1. INTRODUCTION

In the past 50 years acute respiratory distress syndrome (ARDS) remains a significant public health threat with an annual incidence rate of >200,000 cases in the United States and globally with an annual rate of 13.5 cases per 100,000 population [1]. Early detection and trajectory monitoring of ARDS is critical to its treatment [2-5]. Despite the enormous impact on human health, ARDS remains ill-defined, and little progress has been made on advancing novel diagnostic and prognostic tools.

The most common ARDS risk model is Lung Injury Prediction Score (LIPS). However, it has a low ability to predict disease onset [6, 7]. The clinical diagnosis based on the Berlin criteria shows a moderate correlation with real-time and post-mortem tissue pathological findings [8-10], but does not account for the critical, dynamic inflammatory processes that participate in ARDS. Furthermore, presently, there are no predictive clinical biomarkers of ARDS pathogenesis. This hinders the ability of clinicians to detect the early or pre-clinical onset of ARDS.

Breath contains hundreds of volatile organic compounds (VOCs), some of which may provide information about the inflammatory state of the lungs as ARDS develops and progresses. Therefore, breath analysis can potentially be very useful for non-invasive ARDS detection and continuous monitoring [2, 3, 11-14]. In a previous study, we predicted ARDS in human subjects with high accuracy when compared to clinical adjudication based on Berlin criteria [13]. The study was limited by the absence of confirmatory, histopathologic evidence of ARDS (*i.e.*, diffuse alveolar damage). Second, assessment of the rate of change in specific breath compound (*e.g.*, trajectory monitoring), which may be vital for timely clinical and therapeutic decision-making, was limited.

In our present work, we overcame the aforementioned limitations by using a newly developed and optimized the swine model of ARDS that accurately recapitulates the human conditions [15]. Animal models, particularly large animals such as swine, have been used in biomedical research for decades because of their anatomical and physiological similarities to humans [16-18]. They are also an ideal species for modeling the inflammatory response to ARDS [19, 20]. An automated portable GC developed in-house was employed to study swine breath to better understand the pathophysiology and the utility of breath analytes as predictors of ARDS. Each swine was monitored hourly, starting from initiation of the experiment until the development of clinically defined ARDS. This approach provided us the opportunity to monitor VOCs during the

development of ARDS as well as dynamic changes that may be associated with it. Lung pathology was assessed at the end of the experiment to confirm the presence of diffuse alveolar damage.

2. MATERIALS AND METHODS

2.1 Animal Preparation and Induction of Lung Injury

We used a recently-described swine model of ARDS that uses clinically-relevant exposures (sepsis, hyperoxia, volutrauma, and aspiration) and faithfully recapitulates the physiologic, radiographic, and histopathologic features of ARDS [15]. As reported recently, this model fulfills both clinical criteria (Berlin Definition) [9] and experimental criteria for animal models of ARDS [21].

Five male Yorkshire mix swine weighing 45 (±2) kg were used in the study. The Institutional Animal Care and Use Committee (IACUC) at the University of Michigan approved the study. All procedures were carried out in compliance with the university's institutional standards for care and use of laboratory animals in accordance with NIH guidelines [22]. Detailed physiologic data regarding these animals has previously been published [15].

Animals were fasted overnight with libitum access to water and on the day of the experiment, anesthesia was induced. Animals were intubated and mechanically ventilated to maintain a baseline end-tidal CO_2 level between 35-45 mmHg (Biopac Systems Inc., Goleta, CA). At the end of surgical instrumentation, animals were transitioned to total intravenous anesthesia using a combination of propofol, fentanyl, and midazolam. The anesthesia level was monitored for the experiment's duration by assessing corneal reflex, jaw tension, and hemodynamics, including blood pressure, heart rate, and respiratory rate. More details for animal preparation can be found in [15, 23].

Animals were subjected to systemic infection via an injection of E. Coli into the kidney parenchyma (3.5×10^{11} CFU Strain CFT073 E. Coli in a total of approximately 5 mL at a rate of 0.333 mL/min) as previously described [15, 23]. Lung injury was induced utilizing a combination of volume-trauma (tidal volumes of 15 mL/kg), hyperoxia (100% FiO_2), and the instillation of gastric particles (1-2 mL/kg) into the airways via bronchoscope [24]. Following the induction of lung injury, animals were continuously monitored for up to 16 hours at which point the experiment was terminated and the animal was euthanized.

The adjudication of the time points at which ARDS was present was done using the Berlin criteria [9, 10], which uses the ratio of partial pressure of arterial O_2 (PaO_2) to the fraction of inspired oxygen (FiO_2) and the chest radiograph assessed at regular intervals during experiments. The final determination of ARDS was confirmed by lung pathology determined by a pathologist blinded to the specifics of the induced injury. Figures S1 shows the P/F ratios for all five swine animals. Radiographs and pictures of the lungs taken at the end of the experiment are shown elsewhere [15].

2.2. Exhaled Breath Collection and Analysis

We used an automated, portable GC device to sample and analyze the breath from each swine before the induction of lung injury and every hour thereafter for up to 16 hours or until the animal expired, whichever occurred earlier. The GC device was connected to the expiratory port of the ventilator tubing via a T-piece and 1-m long polytetrafluoroethylene (PTFE) tubing (0.64 cm i.d.), as shown in Figure 1. The same GC device was used for breath sampling and analysis for all five animals. For each breath measurement, the breath sample was collected into a thermal desorption tube installed in the GC device at a rate of 70 mL/min, and then analysis was performed. Finally, device cleaning was performed to ensure that no residues were left from the previous run. The details of breath sampling and analysis processes are given in SI 1.2. The total assay time was 30 minutes, including 5 minutes of sample collection at 70 mL/min, 5 minutes of desorption/transfer, 10 minutes of separation, and 10 minutes of cleaning.

The raw chromatograms were first pre-processed for noise reduction, curve smoothing, alignment with the reference chromatogram, and peak assignment. After pre-processing, the area of each peak was calculated and normalized by the entire area under the chromatogram curve [13, 25]. Through machine learning, principal component analysis (PCA), and linear discriminant analysis (LDA) [13, 17, 18], a sub-set of chromatographic peaks were selected as the biomarkers for the discrimination of ARDS. The details about the GC system and the algorithm used in data analysis can be found in [13, 25-27].

Traditional mass spectrometry (MS) was performed to identify the VOCs present in swine breath. The swine's breath at the experiment termination or animal death, whichever occurred first, was collected into a 3 L Tedlar bag. Then it was assayed by our GC coupled to a Thermo Scientific

Single Quadrupole Mass Spectrometer (ISQTM Series) and the results were analyzed using Chromeleon^{TM} 7 Software.

3. RESULTS

Figure 2 shows a representative chromatogram of swine breath. A total of 67 breath chromatograms from the five swine were collected and analyzed, including 34 pre-ARDS chromatograms and 33 chromatograms after development of ARDS (as adjudicated by the Berlin criteria during the experiment and confirmed by lung biopsies at the end of the experiment). There were approximately 60-70 peaks in each individual chromatogram. Collectively, there were a total of 78 different peaks.

3.1. Candidate Biomarkers of ARDS

For biomarker discovery, we used 40 out of 67 breath chromatograms as the training set, among which 20 were pre-ARDS and 20 were ARDS. The 20 pre-ARDS data were from the breath samples taken at the 0th hour (before lung injury was induced), and the 1st, 2nd, and 3rd hour after the induction of lung injury, up until the P/F ratio fell below 300. By doing this, we ensured that the swine had not developed ARDS, which was confirmed by clinical adjudication based on the Berlin criteria. The 20 ARDS chromatograms were from the breath samples taken in the last 4 hours of the experiment for Swine #2, #4, and #5. Swine #1 died within 8 hours and we had only the last two data points adjudicated as ARDS, which are included in the training set. For Swine #3, which died after 16 hours, we used the samples taken in the last 6 hours in the training set. The reason to use the samples taken 4 - 6 hours prior to the animal death is two-fold. First, it is highly likely that the swine had developed ARDS at this stage (4-6 hours prior to death), as confirmed by clinical adjudication. Second, it was necessary to make the total number of ARDS in the training set to be 20 (*i.e.*, equal to the number of pre-ARDS controls). The remaining 27 chromatograms (14 pre-ARDS and 13 ARDS, as adjudicated by the Berlin criteria) were used as the testing set. They were obtained between the 3rd hour after the lung injury and 4-6 hours prior to death.

Through machine learning, PCA, and LDA, seven VOCs (see Table 1) were identified as the candidate biomarker set, which yielded a classification accuracy of 92.5% (Table 2) for the training set and the largest boundary distance in the PCA plot. Interestingly, these seven analytes are the same as seven of the nine biomarkers found in our previous breath analysis that predicted human

ARDS. The similarities in the biomarker set further validate our findings in humans and underscore the clinical relevancy of the swine model to simulate the human conditions. We also examined the other two VOCs, 2,4-dimethylpentane and methylcyclohexane (see Table 1), which differentiated human ARDS, to see whether they can be added to the current the 7-analyte biomarker set. The results were neutral, *i.e*., adding one of them or the combination of both does not change classification accuracy. This may be due to the limited number of swine and the breath samples used in this study. In contrast, the addition of other peak(s) to the current biomarker set resulted in decreased accuracy (Table S1).

In Figure 3, we plotted the normalized peak area for each identified analyte versus time after the induction of lung injury of all five swine to assess the trend in analyte signal during the pathogenesis of ARDS. Of these, 2-dimethyl pentane and 3-methylheptane increased over time, while the other five compounds decreased.

Figure 4 shows the PCA scores for the 40 data points in the training set using the seven discriminating analytes, giving an accuracy of 92.5%. The corresponding statistics is summarized in Table 2. The data demonstrate that pre-ARDS and ARDS are well separated and that the distribution of the pre-ARDS data points is much more clustered than that for the non-ARDS human subjects in our previous study [13], perhaps indicative of the homogeneity of the swine model in comparison with the human subjects. Figures S2 and S3 show the PCA scores for the testing set and for all the data sets, respectively. The corresponding accuracy decreases significantly compared to the training set (see Table 2). However, this does not necessarily indicate that breath analysis fails. Instead, it suggests the limitation of the Berlin criteria, since breath analysis can detect changes consistent with the development of ARDS approximately 3 hours earlier than the clinical adjudication and all of the 27 chromatograms in the testing set were intermediate points obtained during the development of ARDS. See detailed explanation later.

The cross-validation was also performed, where the original training datasets were divided into 5 cross-validation models. The five models yielded an overall classification accuracy of 92.5±1.5% (and 90.5±16.2% for the testing set), which indicates the robustness of the model. The statistics of the five models is presented in Table S3.

3.2. Dynamic Response of Swine Breath

The time trajectory monitoring of disease is critical for guided therapy of ARDS. In our previous study with human subjects, we did not continuously monitor subjects through their entire course so development of an ARDS trajectory was not possible. In contrast, in the current swine experiment we were able to perform a sequential hourly monitoring of the dynamic change of swine ARDS from the very beginning (prior to the induction of lung injury) to the end of the experiment (death of ARDS), as shown in Figure 5 where each swine moved from pre-ARDS to ARDS. The first data point (the 0^{th} hour, before the induction of lung injury) and the last data point (just before the death of ARDS) are well separated and far from the boundary line. For all the five swine, as the experiment progresses, the PCA trajectory moves from the 0^{th} hour (*i.e.*, the pre-ARDS region) towards the boundary line, then reaches the ARDS region, and eventually moves farther away from the boundary line as the severity of ARDS increases. The only exception is "3.10" (Swine #3 at the 10^{th} hour), which shows movement to a decreased state of severity. This could, however, be caused by errors in breath measurement.

4. DISCUSSION AND CONCLUSION

To better understand the physiology, pathophysiology, and therapy for ARDS, we utilized a large animal model that recapitulates the human conditions [15]. The animal model helps overcome some of the current clinical limitations and lends insight into the underlying physiology that leads to ARDS because it utilizes clinically relevant insults leading to diffuse alveolar damage, the histopathologic hallmark of ARDS. To date, most exhaled breath studies of ARDS have been limited in their scope, and most have used a single time point [4]. To our knowledge, we are the first group to use portable GC to study a high fidelity large animal model allowing for continuous point of care (POC) to monitoring of VOCs. The presented results support our previous study based on human subject for delineation between non-ARDS controls and ARDS patients in whom histopathologic examination of lungs are not possible. We found that the seven biomarkers (from swine's breath) as shown in Table 1 are shared with the nine biomarkers (from the human subjects) that could distinguish non-ARDS and ARDS. This underscores the clinical relevancy of the swine model to simulate the human conditions.

The longitudinal analysis revealed that volatile metabolic changes in exhaled breath might be used to map the pathologic trajectory of ARDS. In this animal model we provide evidence that it was possible to detect changes consistent with the development of ARDS 1-6 hours (3 hours on average) earlier than the clinical adjudication based on the Berlin criteria (Table S2). This may be the reason why in the testing set, accuracy is much lower than the training set, because most of the data points in the testing set are obtained when the animals were in the transition from pre-ARDS to ARDS using the Berlin criteria as the gold standard. Also, as expected, animals have different trajectories of ARDS development. For example, it took less than 1 hour for Swine #2 to develop ARDS after the induction of lung injury, whereas it took 3 hours or longer time for the other 4 animals to develop ARDS (they died in 11-16 hours). It should be noted that the ability of breath analysis to provide early detection and trajectory monitoring of ARDS pathogenesis has also been demonstrated in a limited number of human subjects in our previous study (24-48 hours earlier than the ARDS adjudication using the Berlin criteria) [13].

Compared to our previous work on human subjects, the present study includes the following strengths. (1) We were able to control many endogenous and exogenous factors in study subjects that might affect breath VOC profiles. The animals were all male of the same approximate age under the same controlled environmental and dietary conditions. The multiple clinically relevant insults used to produce ARDS were uniform and tightly controlled. The sample collection and analysis were the same for all animals. (2) Lung tissue biopsies were collected at the end of the experiment to examine histopathological changes that are consistent with diffuse alveolar damage to diagnose ARDS, which is not feasible in standard clinical care and clinical research studies. The adjudication of animals' last breath samples compared to the presence of histopathologic adjudicated diffuse alveolar damage provided increased accuracy and confidence of the training set. (3) We were able to perform detailed longitudinal studies (from the healthy pre-ARDS state to the final death due to ARDS), which helps better delineate the trajectory of ARDS. This could be vital for guided therapy and decision making. Finally, we identified a physiologically relevant VOC profile similar to that which we previously found in humans with ARDS [13] (see detailed discussion in Section S3 in the Supplementary Information). As a class of compounds, VOCs are highly relevant to the detection of ARDS because many are alkanes (*e.g.*, pentane, heptane) and alkenes (*e.g.*, α-pinene), which are byproducts of inflammatory processes [3, 28-35]. The found VOC pattern in both our human and swine studies is consistent with what has been previously

reported in the literature [3, 4] and is most often attributable to fatty acid peroxidation [36]. Lipid peroxidation (the oxidative degradation of lipid membranes that leads to cell damage and dysfunction) is mediated by reactive oxygen species, most likely generated by leukocytes [37, 38], which is consistent with what is known about ARDS pathogenesis [39].

The study has several limitations. First, we used only five animals with limited data points. Consequently, we were unable to separate the training and testing set by animal identifiers. As such the training and testing sets have observations from the same animal at different time points. Although we conducted cross-validation, we expect improvement in mean accuracy and standard deviation with increased data sets. Second, severity of ARDS (*i.e.*, mild, moderate, and severe) was not attempted due to lack of transition data points. In future, this work may be possible with increased number of swine and transition data points. Third, we identified seven biomarkers using swine model, compared to nine in the human study. This may be due to the limited number of swine and the breath samples used in this study. Additional swine experiments will be needed to determine the validity of identified compounds for ARDS detection. The institution of therapies that change the trajectory of ARDS by treating underlying causes such as sepsis can be instituted in the future to allow further dynamic perturbation of VOCs to understand whether the technology can track changes prior to standard clinical metrics. Finally, by design our swine model of ARDS is multifactorial in etiology, and our study design does not enable us to determine which exhaled breath signals are attributable to ARDS itself or the constituent exposures of our model (*e.g.*, sepsis, aspiration). Further work will determine the generalizability of our findings.

In conclusion, breath analysis using portable GC has been presented as an alternative for a rapid diagnostic and trajectory monitoring of ARDS. The performance of this methodology in delineating ARDS from pre-ARDS is excellent and provides encouraging conceptual evidence at the experimental level. Continuous monitoring of VOCs produced by ARDS opens up the potential for trajectory monitoring, which will have implications for earlier interventions. Using a high-fidelity swine model to study exhaled breath significantly enhances the ability to study ARDS and to test new technologies aimed at improving its diagnosis, monitoring and treatment.

ACKNOWLEDGMENTS

The authors thank the support from National Institutes of Health (1R21HL139156), Flux HPC Cluster provided by the University of Michigan (UM) Office of Research and Advanced Research Computing–Technology Services, and Analytical Chemistry Lab at the UM Biological Station at Pellston. The development of the swine model of ARDS was supported the Michigan Center for Integrative Research in Critical Care and the Michigan Institute for Clinical and Health Research (MICHR) Accelerating Synergy Award. MICHR is supported by a National Center for Advancing Translational Sciences award (UL1TR002240) from the National Institutes of Health.

REFERENCES

1. Luhr OR, Antonsen K, Karlsson M, Aardal S, Thorsteinsson A, Frostell CG, Bonde J. Incidence and mortality after acute respiratory failure and acute respiratory distress syndrome in Sweden, Denmark, and Iceland. The ARF Study Group. *Am J Respir Crit Care Med* 1999: 159(6): 1849-1861.

2. Bos LD, Schultz MJ, Sterk PJ. Exhaled breath profiling for diagnosing acute respiratory distress syndrome. *BMC Pulm Med* 2014: 14: 72.

3. Bos LD, Weda H, Wang Y, Knobel HH, Nijsen TM, Vink TJ, Zwinderman AH, Sterk PJ, Schultz MJ. Exhaled breath metabolomics as a noninvasive diagnostic tool for acute respiratory distress syndrome. *Eur Respir J* 2014: 44(1): 188-197.

4. Bos LDJ. Diagnosis of acute respiratory distress syndrome by exhaled breath analysis. *Ann Transl Med* 2018: 6(2): 8.

5. Ware LB, Koyama T, Zhao Z, Janz DR, Wickersham N, Bernard GR, May AK, Calfee CS, Matthay MA. Biomarkers of lung epithelial injury and inflammation distinguish severe sepsis patients with acute respiratory distress syndrome. *Critical care (London, England)* 2013: 17(5): R253-259.

6. Gajic O, Dabbagh O, Park PK, Adesanya A, Chang SY, Hou P, Anderson H, 3rd, Hoth JJ, Mikkelsen ME, Gentile NT, Gong MN, Talmor D, Bajwa E, Watkins TR, Festic E, Yilmaz M, Iscimen R, Kaufman DA, Esper AM, Sadikot R, Douglas I, Sevransky J, Malinchoc M. Early identification of patients at risk of acute lung injury: evaluation of lung injury prediction score in a multicenter cohort study. *Am J Respir Crit Care Med* 2011: 183(4): 462-470.

7. Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, Stern EJ, Hudson LD. Incidence and outcomes of acute lung injury. *N Engl J Med* 2005: 353(16): 1685-1693.

8. Kao KC, Hu HC, Chang CH, Hung CY, Chiu LC, Li SH, Lin SW, Chuang LP, Wang CW, Li LF, Chen NH, Yang CT, Huang CC, Tsai YH. Diffuse alveolar damage associated mortality in

selected acute respiratory distress syndrome patients with open lung biopsy. *Crit Care* 2015: 19(1): 228.

9. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, Camporota L, Slutsky AS. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012: 307(23): 2526-2533.

10. Thille AW, Esteban A, Fernández-Segoviano P, Rodriguez JM, Aramburu JA, Peñuelas O, Cortés-Puch I, Cardinal-Fernández P, Lorente JA, Frutos-Vivar F. Comparison of the Berlin definition for acute respiratory distress syndrome with autopsy. *Am J Respir Crit Care Med* 2013: 187(7): 761-767.

11. Carpenter CT, Price PV, Christman BW. Exhaled breath condensate isoprostanes are elevated in patients with acute lung injury or ARDS. *Chest* 1998: 114(6): 1653-1659.

12. Rahimpour E, Khoubnasabjafari M, Jouyban-Gharamaleki V, Jouyban A. Non-volatile compounds in exhaled breath condensate: review of methodological aspects. *Anal Bioanal Chem* 2018: 410(25): 6411-6440.

13. Zhou M, Sharma R, Zhu H, Li Z, Li J, Wang S, Bisco E, Massey J, Pennington A, Sjoding M, Dickson RP, Park P, Hyzy R, Napolitano L, Gillies CE, Ward KR, Fan X. Rapid breath analysis for acute respiratory distress syndrome diagnostics using a portable two-dimensional gas chromatography device. *Anal Bioanal Chem* 2019: 411(24): 6435-6447.

14. Bos LDJ, Walree ICv, 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* 2013: 115(10): 1487-1495.

15. Tiba MH, McCracken BM, Leander DC, Colmenaro CI, Nemzek JA, Sjoding MW, Konopka KE, Flott TL, VanEpps JS, Daniels R, Ward KR, Stringer KA, Dickson RP. A Novel Swine Model of the Acute Respiratory Distress Syndrome Using Clinically-Relevant Injury Exposures. *Physiol Rep* 2021: 9(9): e14871.

16. Swindle MM, Smith AC, Hepburn BJ. Swine as models in experimental surgery. *J Invest Surg* 1988: 1(1): 65-79.

17. Ellis CK, Stahl RS, Nol P, Waters WR, Palmer MV, Rhyan JC, VerCauteren KC, McCollum M, Salman MD. A pilot study exploring the use of breath analysis to differentiate healthy cattle from cattle experimentally infected with Mycobacterium bovis. *PLoS One* 2014: 9(2): e89280-89292.

18. Guamán A, Carreras A, Calvo D, Agudo I, Navajas D, Pardo A, Marco S, Farré R. Rapid detection of sepsis in rats through volatile organic compounds in breath. *J Chromatogr B* 2012: 881-882: 76-82.

19. Katalan S, Falach R, Rosner A, Goldvaser M, Brosh-Nissimov T, Dvir A, Mizrachi A, Goren O, Cohen B, Gal Y, Sapoznikov A, Ehrlich S, Sabo T, Kronman C. A novel swine model of ricin-induced acute respiratory distress syndrome. *Dis Model Mech* 2017: 10(2): 173-183.

20. Matute-Bello G, Frevert CW, Martin TR. Animal models of acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 2008: 295(3): L379-L399.

21. Matute-Bello G, Downey G, Moore BB, Groshong SD, Matthay MA, Slutsky AS, Kuebler WM. An official American Thoracic Society workshop report: features and measurements of experimental acute lung injury in animals. *Am J Respir Cell Mol Biol* 2011: 44(5): 725-738.

22. Guide for the care and use of laboratory animals. *In:* research Ifla, ed. 8th ed. The National Academies press, Washigton D.C; p. 246.

23. Tiba MH, McCracken BM, Dickson RP, Nemzek JA, Colmenero CI, Leander DC, Flott TL, Daniels RC, Konopka KE, VanEpps JS, Stringer KA, Ward KR. A comprehensive assessment of multi-system responses to a renal inoculation of uropathogenic E. coli in swine. *PLOS ONE* 2020: 15(12): e0243577.

24. Nemzek JA, Abatan O, Fry C, Mattar A. Functional contribution of CXCR2 to lung injury after aspiration of acid and gastric particulates. *Am J Physiol Lung Cell Mol Physiol* 2010: 298(3): L382-L391.

25. Sharma R, Zhou M, Hunter MD, Fan X. Rapid In Situ Analysis of Plant Emission for Disease Diagnosis Using a Portable Gas Chromatography Device. *J Agric Food Chem* 2019: 67(26): 7530-7537.

26. Gillies CE, Jennaro TS, Puskarich MA, Sharma R, Ward KR, Fan X, Jones AE, Stringer KA. A Multilevel Bayesian Approach to Improve Effect Size Estimation in Regression Modeling of Metabolomics Data Utilizing Imputation with Uncertainty. *Metabolites* 2020: 10(8): 319.

27. Sharma R, Zang W, Zhou M, Schafer N, Begley LA, Huang YJ, Fan X. Real Time Breath Analysis Using Portable Gas Chromatography for Adult Asthma Phenotypes. *Metabolites* 2021: 11(5): 265.

28. Alkhouri N, Singh T, Alsabbagh E, Guirguis J, Chami T, Hanouneh I, Grove D, Lopez R, Dweik R. Isoprene in the Exhaled Breath is a Novel Biomarker for Advanced Fibrosis in Patients with Chronic Liver Disease: A Pilot Study. *Clin Transl Gastroenterol* 2015: 6(9): 1-7.

29. Miekisch W, Schubert JK, Noeldge-Schomburg GFE. Diagnostic potential of breath analysis-focus on volatile organic compounds. *Clin Chim Acta* 2004: 347: 25-39.

30. Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, O'Neill HJ. Volatile organic compounds in exhaled air from patients with lung cancer. *Clin Chem* 1985: 31(8): 1278-1282.

31. O'Neill HJ, Gordon SM, O'Neill MH, Gibbons RD, Szidon JP. A computerized classification technique for screening for the presence of breath biomarkers in lung cancer. *Clin Chem* 1988: 34(8): 1613-1618.

32. Sharma JN, Al-Omran A, Parvathy SS. Role of nitric oxide in inflammatory diseases. *Inflammopharmacology* 2007: 15: 252-259.

33. Bhatia M. H_2S and Inflammation: An Overview. *Handb Exp Pharmacol* 2015: 230: 165-180.

34. Wallace JL, Ferraz JGP, Muscara MN. Hydrogen Sulfide: An Endogenous Mediator of Resolution of Inflammation and Injury. *Antioxid Redox Signal* 2012: 17: 58-67.

35. Ialenti A, Ianaro A, Moncada S, Rosa M. Modulation of acute inflammation by endogenous nitric oxide. *Eur J Pharmacol* 1992: 211: 177-182.

36. Yin H, Xu L, Porter NA. Free radical lipid peroxidation: mechanisms and analysis. *Chem Rev* 2011: 111(10): 5944-5972.

37. Kneepkens CM, Ferreira C, Lepage G, Roy CC. The hydrocarbon breath test in the study of lipid peroxidation: principles and practice. *Clin Invest Med* 1992: 15(2): 163-186.

38. Riely CA, Cohen G, Lieberman M. Ethane Evolution: A New Index of Lipid Peroxidation. *Science* 1974: 183(4121): 208-210.

39. Chabot F, Mitchell JA, Gutteridge JM, Evans TW. Reactive oxygen species in acute lung injury. *Eur Respir J* 1998: 11(3): 745-757.

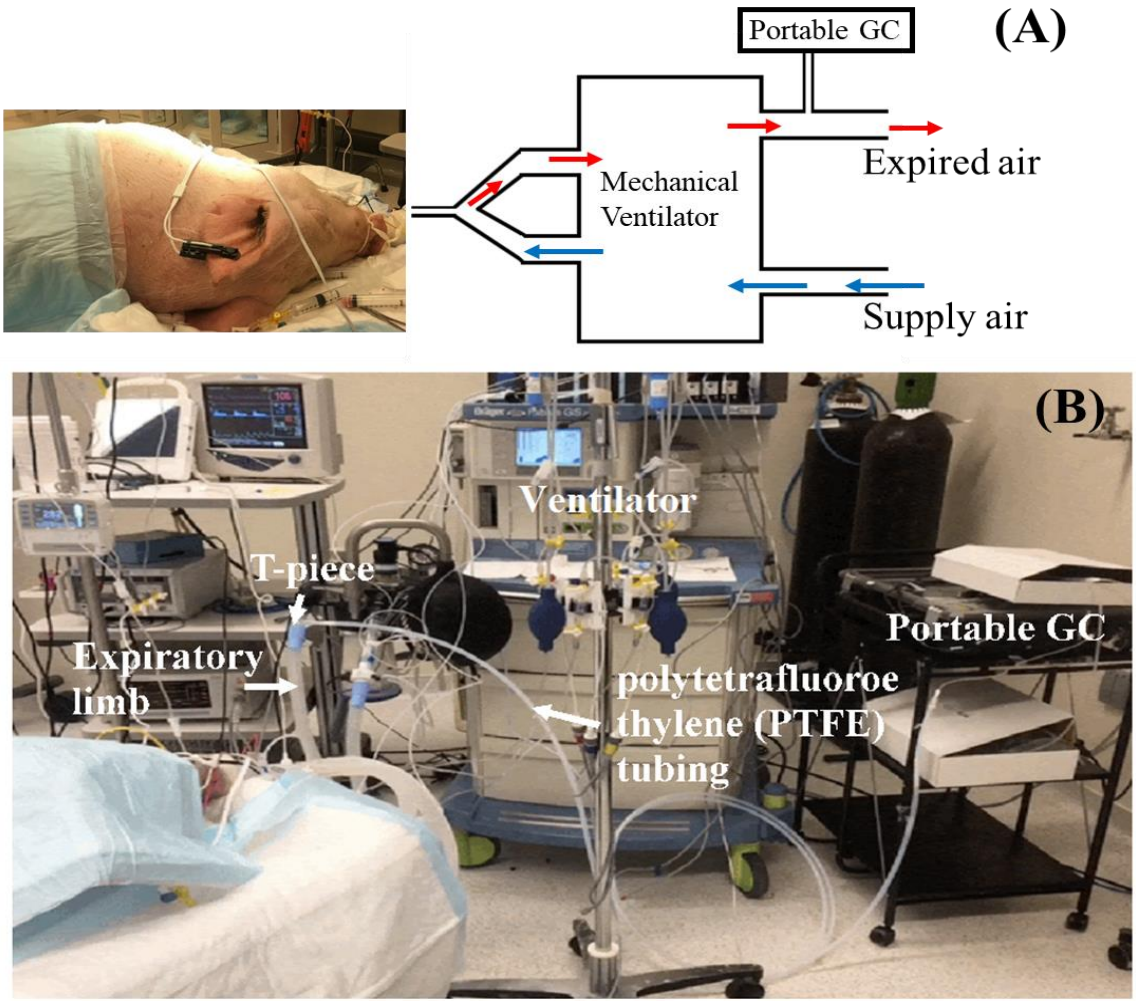

Figure 1. (A) Schematic of a portable GC attached to a ventilator for breath analysis. (B) Picture taken during the breath measurement, showing the connection of the portable GC to a mechanical ventilator via a sampling tube to monitor exhaled breath of a swine.

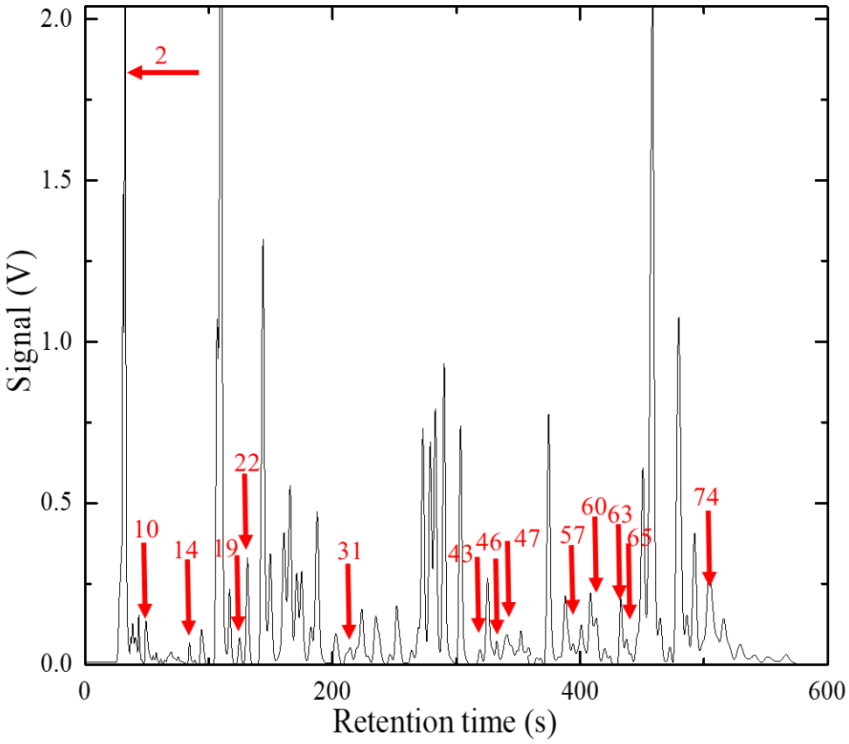

Figure 2. Representative GC chromatogram of exhaled breath from a swine animal via a ventilator. The red arrows show the locations of all peaks used in biomarker search (see details in Table S1).

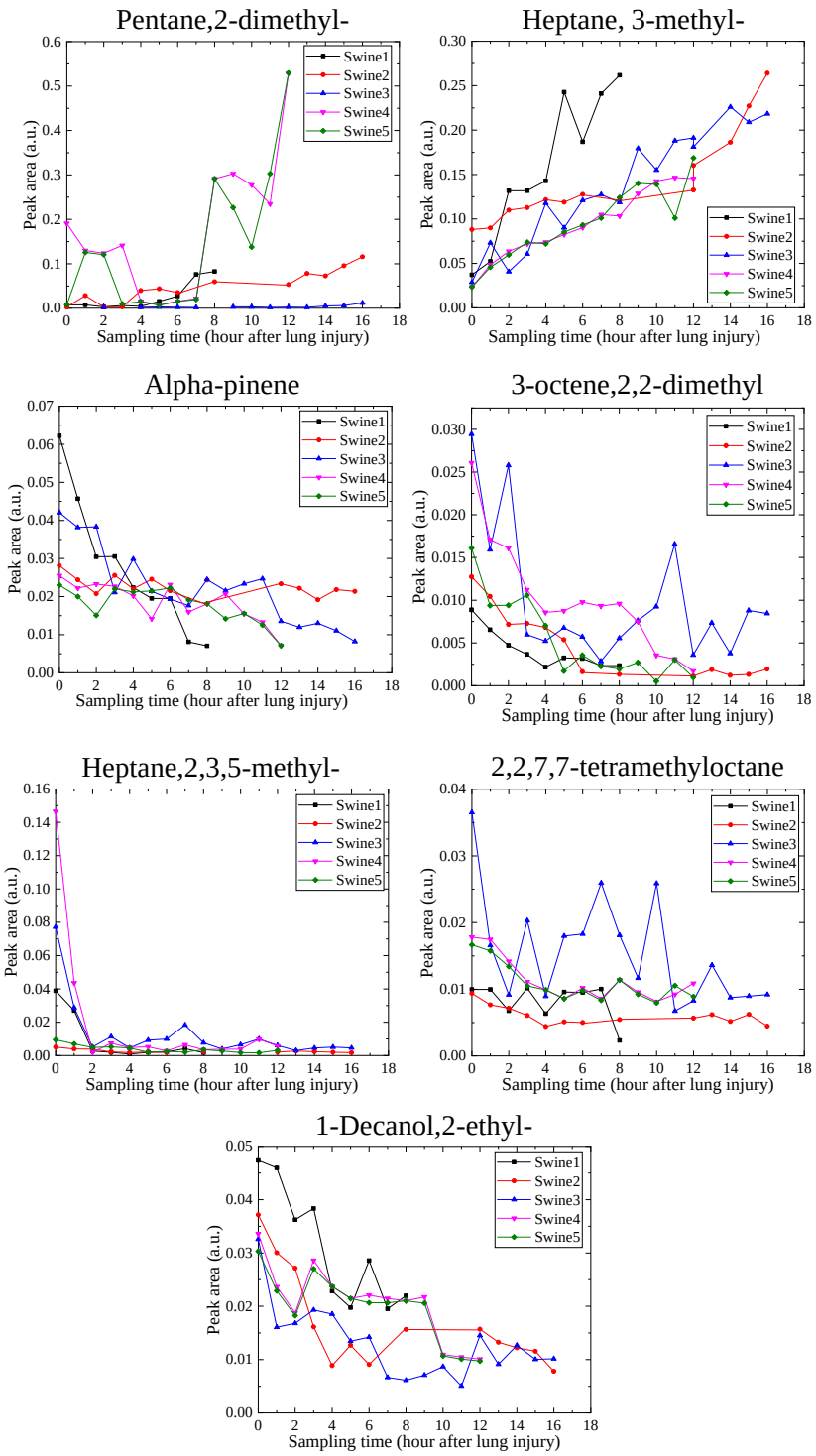

Figure 3. Normalized peak area vs. time of seven biomarkers after the induction of lung injury. Each peak area is normalized to the total area under the chromatogram curve. The 0^{th} hour refers to the time just before the induction of lung injury.

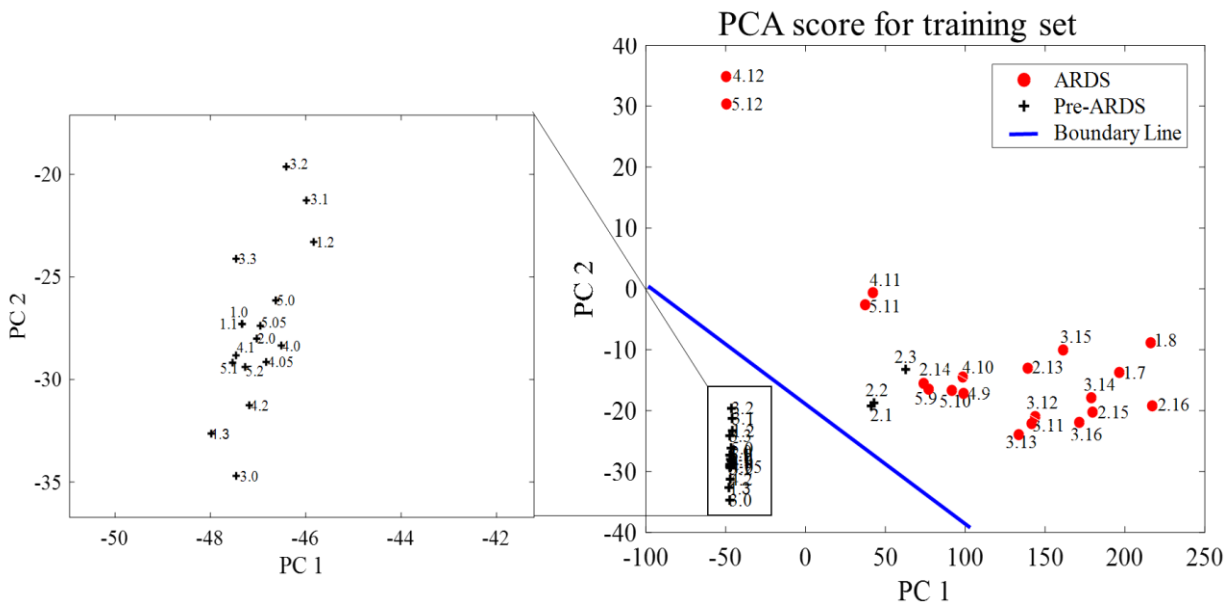

Figure 4. PCA plot for the training set (40 data points in total). Black crosses and red circles denote respectively Pre-ARDS and ARDS adjudicated using the Berlin criteria. The region below and above the boundary line represents the baseline and ARDS region, respectively, as determined by breath analysis using the seven biomarkers in Table 1. The numbers shown in the PCA plot denote "swine number.sampling time". For example: "1.2" refers to Swine #1 sampled 2 hours after the induction of lung injury. "4.0" refers to Swine #4 sampled before the induction of lung injury. "4.11" refers to Swine #4 sampled 11 hours after the induction of lung injury.

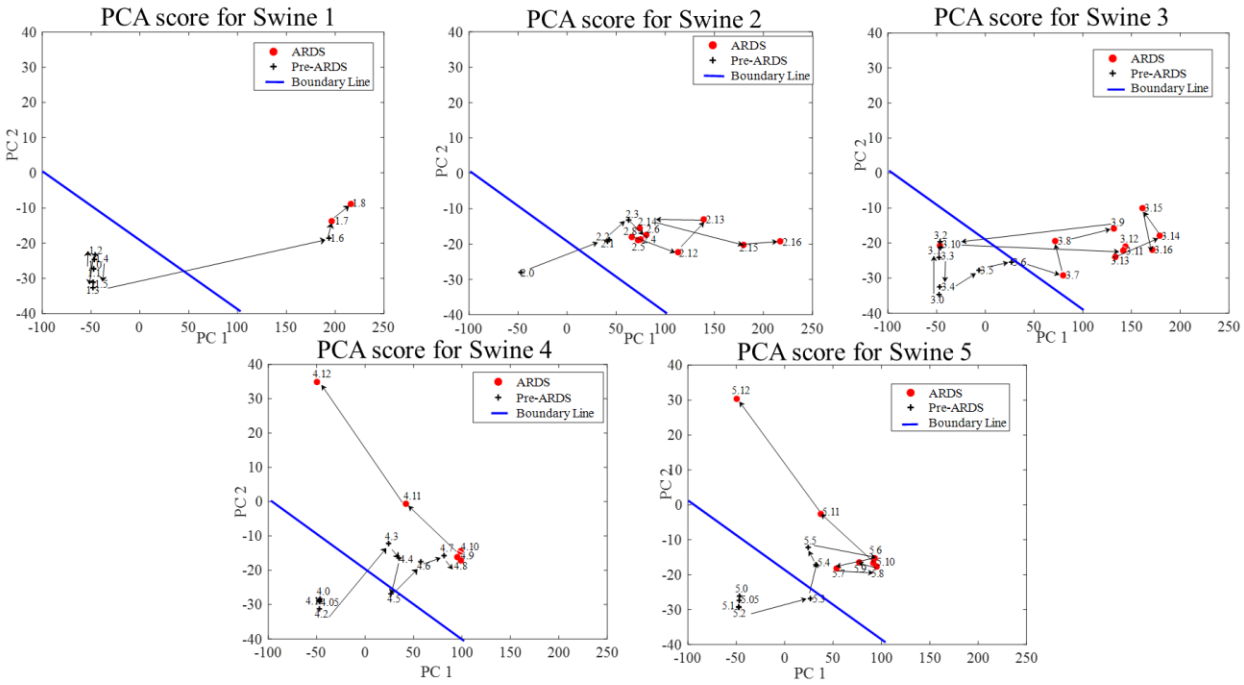

Figure 5. PCA plot showing hourly trajectory of individual swine starting from the 0th hour (healthy and just prior to the induction of lung injury) to the end of the experiment or until the animal died (last data point, ARDS confirmed by biopsy). This figure shows the dynamic change in swine's breath when the animal status changes from healthy pre-ARDS to ARDS. Black crosses and red circles denote respectively pre-ARDS and ARDS adjudicated based on the Berlin criteria.

Peak ID	Chemical Name	CAS Number	Formula
2	Pentane, 2-methyl-	107-83-5	C_6H_{14}
22	Heptane, 3-methyl-	589-81-1	C_8H_{18}
63	2,2,7,7-Tetramethyloctane	1071-31-4	$C_{12}H_{26}$
65	1-Decanol, 2-ethyl-	21078-65-9	$C_{12}H_{26}O$
47	3-Octene, 2,2-dimethyl-	86869-76-3	$C_{10}H_{20}$
46	α-Pinene	80-56-8	$C_{10}H_{16}$
57	Heptane, 2,3,5-trimethyl-	20278-85-7	$C_{10}H_{22}$
10	*Pentane, 2,4-dimethyl-*	*108-08-7*	C_7H_{16}
14	*Cyclohexane, methyl-*	*108-87-2*	C_7H_{14}

Table 1. Seven breath biomarkers that distinguish pre-ARDS and ARDS (the top seven). The last two VOCs (2,4-dimethyl-pentane and methyl-cyclohexane) were used as additional biomarkers to distinguish non-ARDS and ARDS in our previous studies with human subjects.

	Training statistics			Testing statistics			Overall statistics		
	ARDS	**Pre-ARDS**	**Total**	**ARDS**	**Pre-ARDS**	**Total**	**ARDS**	**Pre-ARDS**	**Total**
Positive (ARDS)	20	3	23	12	7	19	32	10	42
Negative (Pre-ARDS)	0	17	17	1	7	8	1	24	25
Column total	20	20	40	13	14	27	33	34	67
Specificity (%)	85%			50%			70.6%		
Sensitivity (%)	100%			92.3%			97%		
Positive predictive value (%)	87%			63%			76.2%		
Negative predictive value (%)	100%			87.5%			96%		
Total accuracy	92.5%			70.3%			83.6%		

Table 2. Corresponding statistics based on the PCA scores of breath analysis for training, test, and overall sets.

Supplementary Information for

Breath Analysis for Detection and Trajectory Monitoring of Acute Respiratory Distress Syndrome in Swine

Ruchi Sharma[1], Menglian Zhou[1], Mohamad Hakam Tiba[2,6], Brendan M. McCracken[2,6], Robert P. Dickson[3,6], Christopher E. Gillies [2,6,8], Michael W. Sjoding[3,6], Jean A. Nemzek[4,5,6], Kevin R. Ward[1,2,6], Kathleen A. Stringer[3,6,7], and Xudong Fan[1,6,+]

[1]Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, U.S.A.

[2]Department of Emergency Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[3]Department of Internal Medicine, Division of Pulmonary and Critical Care Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[4]Unit of Laboratory Animal Medicine, University of Michigan, Ann Arbor, MI, U.S.A.

[5]Department of Pathology, University of Michigan, Ann Arbor, MI, U.S.A.

[6]Michigan Center for Integrative Research in Critical Care, University of Michigan, Ann Arbor, MI, U.S.A.

[7]Department of Clinical Pharmacy, College of Pharmacy, University of Michigan, Ann Arbor, MI, U.S.A

[8]Michigan Institute for Data Science (MIDAS), Office of Research, University of Michigan, Ann Arbor, MI, U.S.A

[+]xsfan@umich.edu

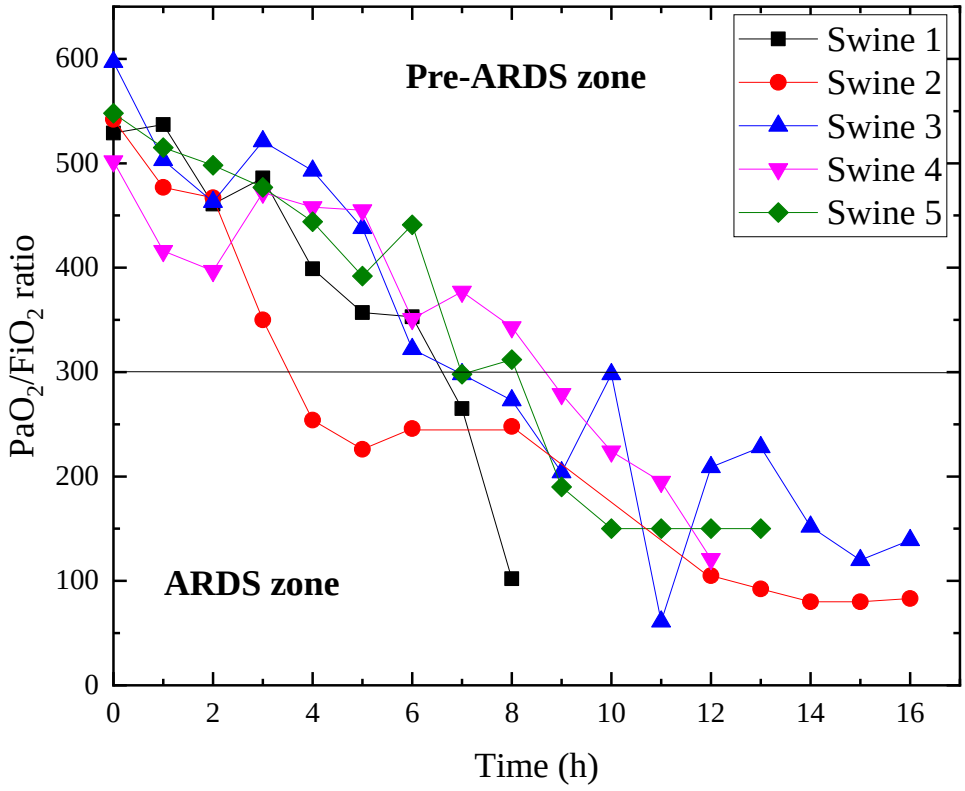

Figure S1. The PaO_2/FiO_2 (P/F) ratios for all five swine. The arterial blood samples were collected and tested every hour, starting from the baseline (the 0^{th} hour) to the end of the experiment. According to the Berlin criteria, the zone below/above the boundary line represents the ARDS/pre-ARDS region, respectively.

S1. Selection of biomarkers

In our study, a total of m=78 peaks were found. We first assumed that there are n=4 peaks relevant to the classification of ARDS versus pre-ARDS. We found that the 4-peak subset of (2, 22, 63, 65) (note: the number in the parentheses is the Peak ID of a peak in the chromatogram) provides the best classification with an accuracy of 75% (see Table S1). Then additional 5 peaks (n'=5) were added, and we found that the 9-peak subset of [(2, 22, 63, 65) + (47, 46, 57, 10, 14)] provides the best classification with an accuracy of 92.5%. Then another 5 peaks (n'=5) were added and we found that the 14-peak subset of [(2, 22, 63, 65) + (47, 46, 57, 10, 14) + (19, 31, 43, 60, 74)] provides the best classification with an accuracy of 90%. The classification accuracy becomes worse from the 9-peak subset to the 14-peak subset. Therefore, we discarded the 14-peak subset and focused only on the 9 peaks [(2, 22, 63, 65) + (47, 46, 57, 10, 14)]. We found that the 7-peak subset of [(2, 22, 63, 65) + (47, 46, 57)] provides the same classification with a total accuracy of 92.5% as that of 9-peak subset. Peak 10, Peak 14, or the combination of both, did not yield higher accuracy. Consequently, we chose the 7 peaks [(2, 22, 63, 65) + (47, 46, 57)] as the biomarkers to distinguish pre-ARDS and ARDS.

	Peak ID			
	4-Peak search	**7-Peak search**	**9-Peak search**	**14-Peak search**
	2	2	2	2
	22	22	22	22
	63	63	63	63
	65	65	65	65
		47	47	47
		46	46	46
		57	57	57
			10	10
			14	14
				19
				31
				43
				60
				74
Accuracy (%)	75	92.5	92.5	90

Table S1: Accuracy tabulated for all peak search for peak subset containing 4 peaks, 7 peaks, 9 peaks, and 14 peaks for the training set. The same color indicates the same subset of peaks kept while doing the additional peak search.

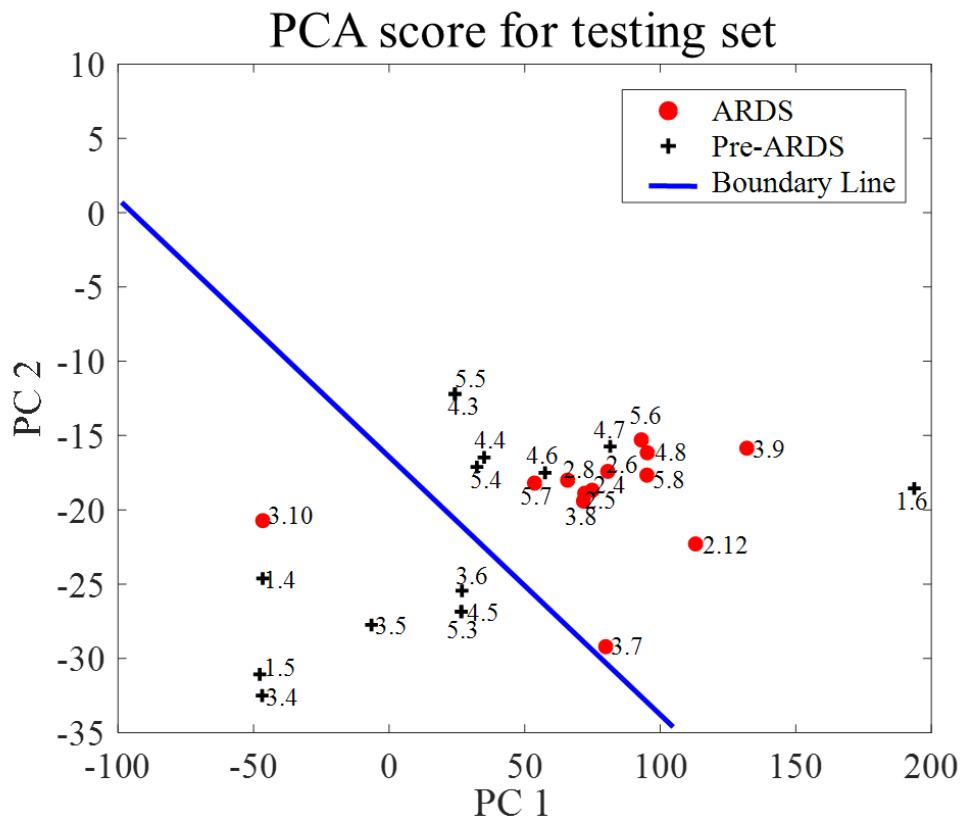

Figure S2. PCA plot of the testing set (27 data points). The region below and above the boundary line represents pre-ARDS and ARDS, respectively.

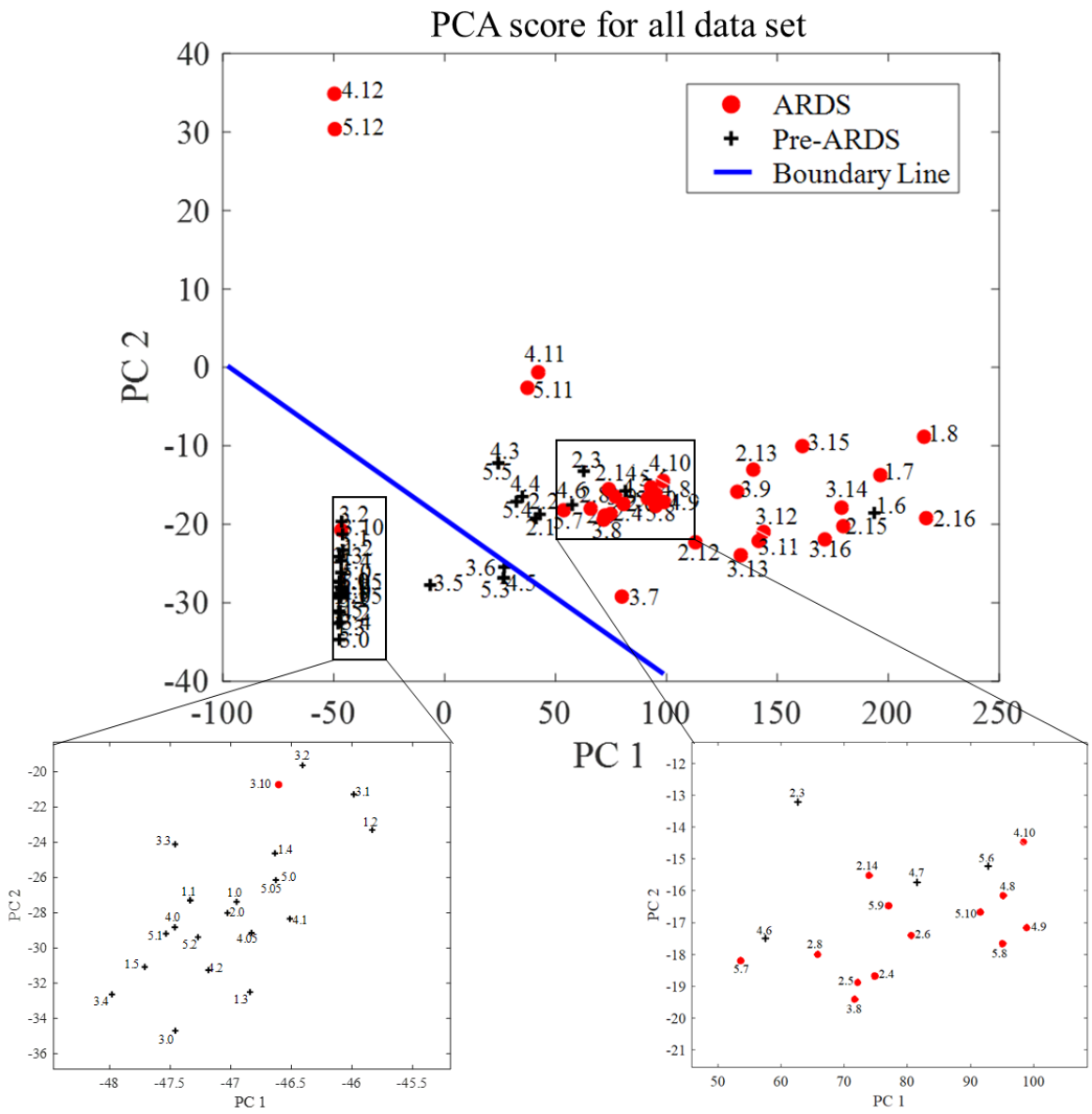

Figure S3. PCA plot of all data points (training and testing sets combined).

Swine	**#1**	**#2**	**#3**	**#4**	**#5**
Berlin criteria	7^{th} hour	4^{th} hour	7^{th} hour	9^{th} hour	7^{th} hour
Breath	6^{th} hour	1^{st} hour	7^{th} hour	3^{rd} hour	4^{th} hour

Table S2. Comparison of ARDS onset detection between the Berlin criteria and breath analysis. The data are extracted from Figure 5 in the main text.

	Training statistics				
Cross Validation	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Specificity (%)	100	100	81.2	81.2	81.2
Sensitivity (%)	100	100	100	100	100
Positive predictive value (%)	100	100	82.4	84.2	84.2
Negative predictive value (%)	100	100	100	100	100
Total accuracy (%)	100	100	90	90.6	90.6
	Testing statistics				
Cross Validation	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Specificity (%)	100	25	75	100	100
Sensitivity (%)	100	100	100	100	100
Positive predictive value (%)	100	57.1	85.7	100	100
Negative predictive value (%)	100	100	100	100	100
Total accuracy (%)	100	62.5	90	100	100
	Overall statistics				
Cross Validation	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Specificity (%)	88.9	85	80	85	85
Sensitivity (%)	100	100	100	100	100
Positive predictive value (%)	90.9	86.95	83.3	86.95	86.95
Negative predictive value (%)	100	100	100	100	100
Total accuracy (%)	95	92.5	90	92.5	92.5

Table S3. Statistics for cross-validation. The original training datasets were divided into 5 cross-validation models. In Model 1, Swines #2-5 datasets were used as the training set and Swine #1 dataset as the testing set. In Model 2, Swines #1 and 3-5 datasets were used as the training set and Swine #2 dataset as the testing set, and so on so forth. Note that the accuracy for the test set (Swine #2) decreases significantly in Model 2 since Swine #2 might have developed ARDS the 1st hour after the induction of lung injury (see "PCA score for Swine 2" in Figure 5), although the clinical adjudication based on the Berlin criteria suggested otherwise.

S2. Operation of the portable GC

The operation procedures and parameters of the portable GC are described as follows.

(1) Sampling: Breath VOCs were drawn by the diaphragm pump through the 2-port valve and adsorbed by the thermal desorption tube at a flow rate of 70 mL/min for 5 min (a total volume of 350 mL). The optimization of the flow rate resulted from a balance between reducing sampling time and preventing VOC breakthrough in the thermal desorption tube. The sample volume (350 mL) was optimized to achieve adequate signal-to-noise ratios for most VOC peaks while not saturating the detector.

(2) Desorption and injection: The 2-port valve was closed and helium gas was flowed through the 3-port valve to provide the carrier gas at a flow rate of 2 mL/min. Meanwhile, the thermal desorption tube was heated to 300 °C for 5 min to transfer the trapped analytes onto the micro-thermal injector. Then the micro-thermal injector was heated to 250 °C in 0.3 s and then kept at 250 °C for 5 s for complete thermal desorption and injection of the analytes into the column. The micro-thermal injector heating parameter was optimized to desorb all VOCs and achieve sharp injection peak width (~0.5 s full-width-at-half-maximum).

(3) Separation: The analytes underwent separation through the 10 m long column and were then detected by the μPID. During the separation, the column was kept at 25 °C for 2 min, then first ramped at a rate of 10 °C min^{-1} to 80 °C, next ramped at a rate of 40 °C min^{-1} to 120 °C, and kept at 120 °C for 1 min. The helium flow rate was 2 mL/min for the column. The ramp rate, column temperature, and carrier gas flow rate were optimized to achieve the best separation of breath VOCs with the shortest possible time.

(4) Cleaning: After analysis, the thermal desorption tube was heated to 300 ºC for 5 min followed by heating the micro-thermal injector to 250 °C in 0.3 s and then keeping it at 250 °C for 6 s at a helium flow rate of 25 mL/min. This process was repeated twice in order to completely remove residual analytes (if any) trapped in the thermal desorption tube and the micro-thermal injector.

The total assay time was 30 minutes, which included 5 minutes of sample collection, 5 minutes of desorption/transfer, 10 minutes of separation, and 10 minutes of cleaning.

S3. Comparison of biomarkers between swine and human

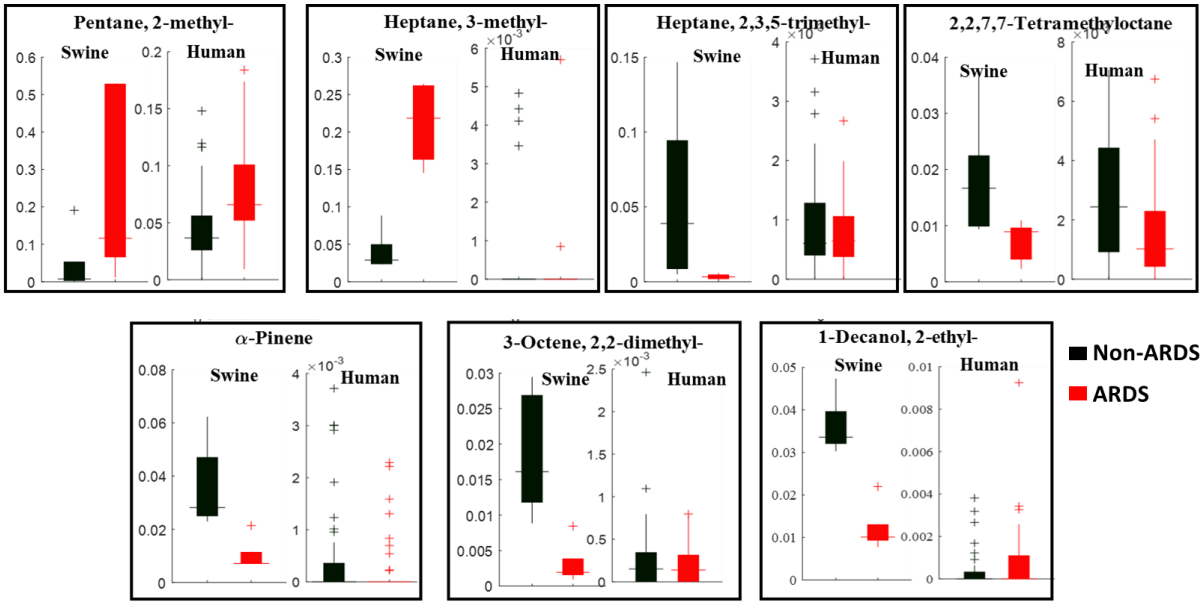

Figure S4. Comparison of the same seven ARDS biomarkers between swine and human.

Figure S4 compares the same seven ARDS biomarkers that we found in the current study (swine) and our previous human study [1]. The black boxplots are for pre-ARDS or non-ARDS, whereas the red boxplots are for ARDS. The swine's data are extracted from Figure 3 in the main text. For each biomarker, we use the first five datapoints (one for each animal) collected at the very beginning of the experiment (time = 0 hour, prior to the induction of lung injury) and last five datapoints (one for each animal) collected at the end of our time-series breath measurement. The human's data are extracted from our previous study [1], in which there were 44 non-ARDS samples and 41 ARDS samples. It should be emphasized that based on our algorithm the entire seven biomarkers must be used as a whole set. Comparing of each individual markers may result in a misleading conclusion.

With the above explanation, let us compare those seven biomarkers. In general, for each biomarker the difference between non-ARDS (or pre-ARDS) and ARDS in swine is much more distinct than in human. Among the seven biomarkers, the trend of five biomarkers (2-methylpentane, 2,3,5-trimethylheptane, 2,2,7,7-tetramethylocatane, α-pinene, and 2,2,-dimethyl-3-octene) match well between swine and human. The remaining two biomarkers, 3-methylhepane and 2-ethyl-1-decanol, change signifiantly from pre-ARDS to ARDS in the swine model, whereas in human the changes are quite small. The strong overlap of each biomarker between non-ARDS and ARDS in human reflects the heterogeneity of human subjects (as well as clinical intervention on the human subjects during treatment) as compared to the homogeneous swine model (without any intervention), which we have also dissused in the main text.

Finally, we need to add the following two comments. First, despite the small difference of 3-methyl-hepane and 2-ethyl-1-decanol between non-ARDS and ARDS, these two compounds need to be included in the whole biomarker set. Without them, the accuracy of human ARDS breath analysis would decrease. Second, the PCA coefficients that we used on the seven (out of nine) biomarkers in the human study cannot be directly applied to the same seven biomarkers in the

swine model in order to determine the status of swine (*e.g.*, pre-ARDS or ARDS), since the compounds in human and swine breaths are not exactly the same.

[1] Zhou M, Sharma R, Zhu H, Li Z, Li J, Wang S, Bisco E, Massey J, Pennington A, Sjoding M, Dickson RP, Park P, Hyzy R, Napolitano L, Gillies CE, Ward KR, Fan X. Rapid breath analysis for acute respiratory distress syndrome diagnostics using a portable two-dimensional gas chromatography device. *Anal Bioanal Chem* 2019: 411(24): 6435-6447.