

## **Urine-based multi-omic comparative analysis of COVID-19 and bacterial sepsis-induced ARDS**

Richa Batra<sup>1\*</sup>, Rie Uni<sup>2\*</sup>, Oleh M. Akchurin<sup>3,4</sup>, Sergio Alvarez-Mulett<sup>5</sup>, Luis G. Gómez-Escobar<sup>5</sup>, Edwin Patino<sup>2</sup>, Katherine L. Hoffman<sup>6</sup>, Will Simmons<sup>6</sup>, Kelsey Chetnik<sup>1</sup>, Mustafa Buyukozkan<sup>1</sup>, Elisa Benedetti<sup>1</sup>, Karsten Suhre<sup>7</sup>, Edward Schenck<sup>5</sup>, Soo Jung Cho<sup>5</sup>, Augustine M.K. Choi<sup>5</sup>, Frank Schmidt<sup>8#</sup>, Mary E. Choi<sup>2#</sup>, Jan Krumsiek<sup>1#</sup>

<sup>1</sup>Department of Physiology and Biophysics, Institute for Computational Biomedicine, Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY 10021, USA

<sup>2</sup>Division of Nephrology and Hypertension, Joan and Sanford I. Weill Department of Medicine, New York, NY, USA

<sup>3</sup>Department of Pediatrics, Division of Pediatric Nephrology, Weill Cornell Medicine, New York, NY, USA

<sup>4</sup>New York-Presbyterian Hospital, New York, NY, USA

<sup>5</sup>Department of Medicine, Division of Pulmonary and Critical Care Medicine, Weill Cornell Medicine, New York, NY, USA

<sup>6</sup>Department of Population Health Sciences, Division of Biostatistics, Weill Cornell Medicine, New York, NY, USA

<sup>7</sup>Bioinformatics Core, Weill Cornell Medicine – Qatar, Qatar Foundation, Doha, Qatar

<sup>8</sup>Proteomics Core, Weill Cornell Medicine – Qatar, Qatar Foundation, Doha, Qatar

\*joint first authors, #corresponding authors

## Abstract

Acute respiratory distress syndrome (ARDS), a life-threatening condition during critical illness, is a common complication of COVID-19. It can originate from various disease etiologies, including severe infections, major injury, or inhalation of irritants. ARDS poses substantial clinical challenges due to a lack of etiology-specific therapies, multisystem involvement, and heterogeneous, poor patient outcomes. A molecular comparison of ARDS groups holds the potential to reveal common and distinct mechanisms underlying ARDS pathogenesis. In this study, we performed a comparative analysis of urine-based metabolomics and proteomics profiles from COVID-19 ARDS patients (n = 42) and bacterial sepsis-induced ARDS patients (n = 17). The comparison of these ARDS etiologies identified 150 metabolites and 70 proteins that were differentially abundant between the two groups. Based on these findings, we interrogated the interplay of cell adhesion/extracellular matrix molecules, inflammation, and mitochondrial dysfunction in ARDS pathogenesis through a multi-omic network approach. Moreover, we identified a proteomic signature associated with mortality in COVID-19 ARDS patients, which contained several proteins that had previously been implicated in clinical manifestations frequently linked with ARDS pathogenesis. In summary, our results provide evidence for significant molecular differences in ARDS patients from different etiologies and a potential synergy of extracellular matrix molecules, inflammation, and mitochondrial dysfunction in ARDS pathogenesis. The proteomic mortality signature should be further investigated in future studies to develop prediction models for COVID-19 patient outcomes.

## 1. Introduction

The ongoing SARS-CoV-2 induced coronavirus disease 2019 (COVID-19) pandemic has been a major impediment to human life globally<sup>1,2</sup>. One of the main complications of severe COVID-19 is acute respiratory distress syndrome (ARDS). ARDS is a common presentation of critical illnesses, including severe infections, major injury, or inhalation of irritants<sup>3</sup>. While COVID-19-related ARDS and ARDS originating from other pathologies (hereby referred to as non-COVID-19 ARDS) have overlapping clinical features, COVID-19 ARDS is characterized by a protracted hyperinflammatory state and higher rates of thrombosis<sup>4-13</sup>. The field currently lacks etiology-specific therapies and reliable predictors of heterogeneous patient outcomes<sup>14</sup>.

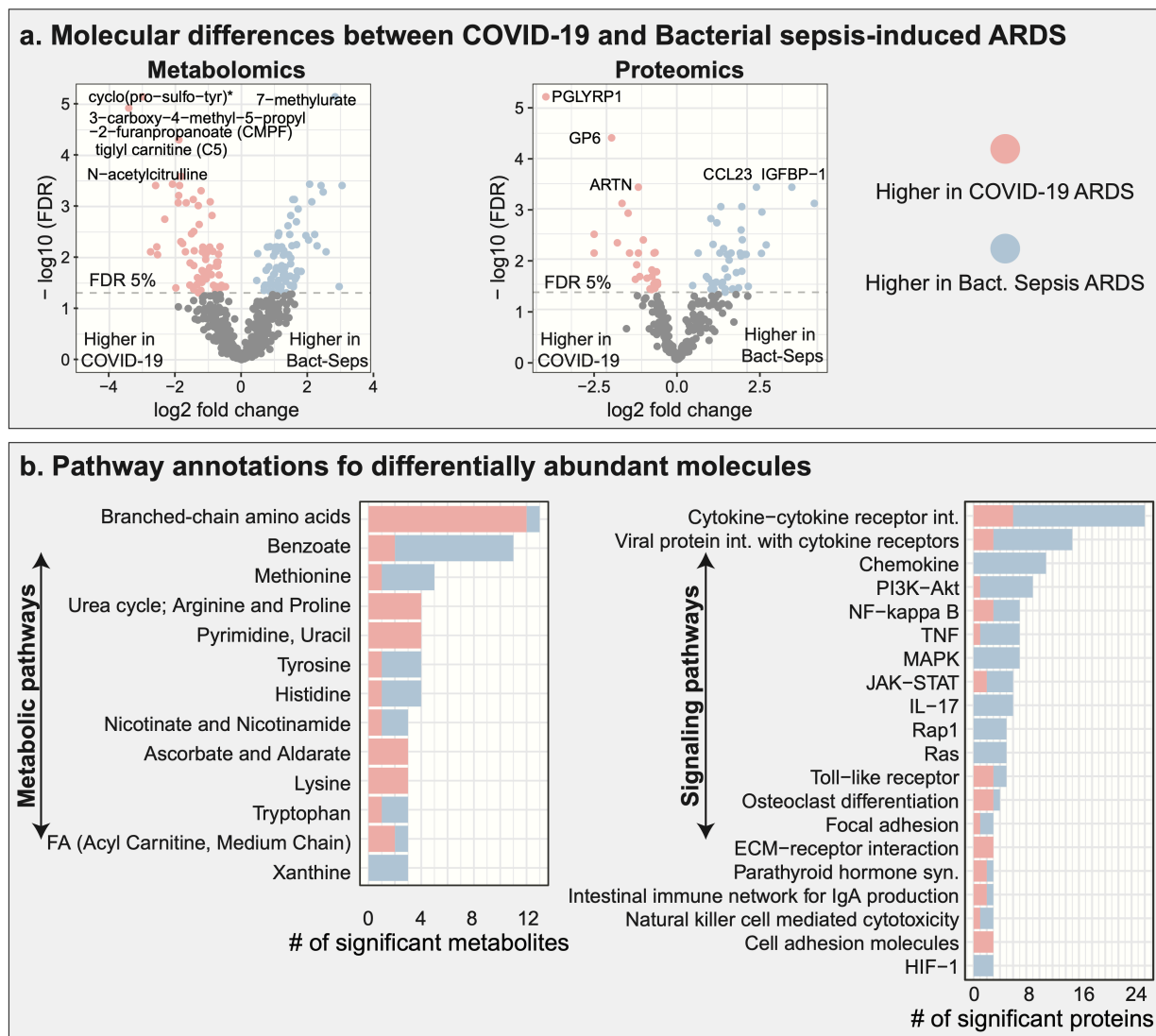
To address these critical knowledge gaps, we recently elucidated molecular differences between and within two ARDS etiologies - COVID-19 and bacterial sepsis<sup>15</sup>. Extending this blood-based ARDS comparison, we here performed a similar analysis on urine samples. It has been suggested that urine-based molecular profiles reflect an individual's physiological changes<sup>16</sup> and have the potential to be used as diagnostic and prognostic biomarkers<sup>17-20</sup>. Previous urine-based COVID-19 studies have made substantial efforts to determine molecular markers distinguishing COVID-19 from healthy controls or less severe COVID-19 cases<sup>21-24</sup>. However, a detailed comparison of the molecular differences between two ARDS groups has so far been missing.

In this study, we analyzed urine samples from 59 ARDS patients, with COVID-19 (n = 42) and bacterial sepsis diagnosis (n = 17). We followed a two-step analysis workflow to elucidate the differences between the two ARDS groups. In the first part, we compared metabolomic and proteomic profiles between the two groups to identify differentially abundant molecules. For a systematic cross-omics analysis of these molecules, we performed a data-driven network analysis. In the second part of the study, we compared the molecular heterogeneity within each ARDS group. To this end, we associated the omics measurements with clinical manifestations, including acute kidney injury (AKI) incidence, platelet counts, PaO<sub>2</sub>/FiO<sub>2</sub>, and mortality. For further exploration and reproducibility of our findings, we share all results, analysis scripts, and de-identified omics data.

## 2. Results and Discussion

### 2.1. Molecular associations differentiating COVID-19 and bacterial sepsis-induced ARDS

To identify the molecular differences between COVID-19 ARDS and bacterial sepsis-induced ARDS, urine-based metabolomic and proteomic profiles from 59 samples were analyzed (n = 42 COVID-19, and n = 17 bacterial sepsis). At a 5% false discovery rate (FDR), 220 molecules were significantly different between the two groups, representing 150 metabolites (70 higher in COVID-19 ARDS and 80 lower), and 70 proteins (28 higher in COVID-19 and 42 lower) (**Figure 1a**). The results of this analysis are available in **Supplementary Table 1**. To aid the functional interpretation of these molecules, metabolites and proteins were annotated with 'sub-pathway' annotations provided by Metabolon and proteins were annotated with KEGG pathways<sup>25</sup> (**Supplementary Table 2**). Top ranking pathways are shown in **Figure 1b**. Two of the pathways we identified in this ARDS comparison, extracellular matrix (ECM) and cell adhesion molecules (CAMs), have also been implicated in previous urine-based studies comparing COVID-19 with a control group<sup>26</sup>. In addition, blood-based studies have reported several of these pathways in the context of COVID-19 ARDS when compared to healthy controls, including amino acid metabolism, lipid metabolism, urea cycle, MAPK, PI3K-Akt, and JAK-STAT signaling<sup>27-31</sup>. Taken together, we identified 220 molecules that were differentially abundant between the two ARDS groups, with 33 distinct biological pathways that had three or more significant molecules.



**Figure 1: Molecular signature of COVID-19 ARDS compared to bacterial sepsis-induced ARDS.** **a.** Differentially abundant molecules (150 metabolites, 70 proteins) between the two ARDS groups. **b.** Functional annotations of differentially abundant metabolites and proteins at the pathway level. Overall, 33 metabolic and signaling pathways with three or more significant molecules were deregulated between the two ARDS groups. FA=fatty acid.

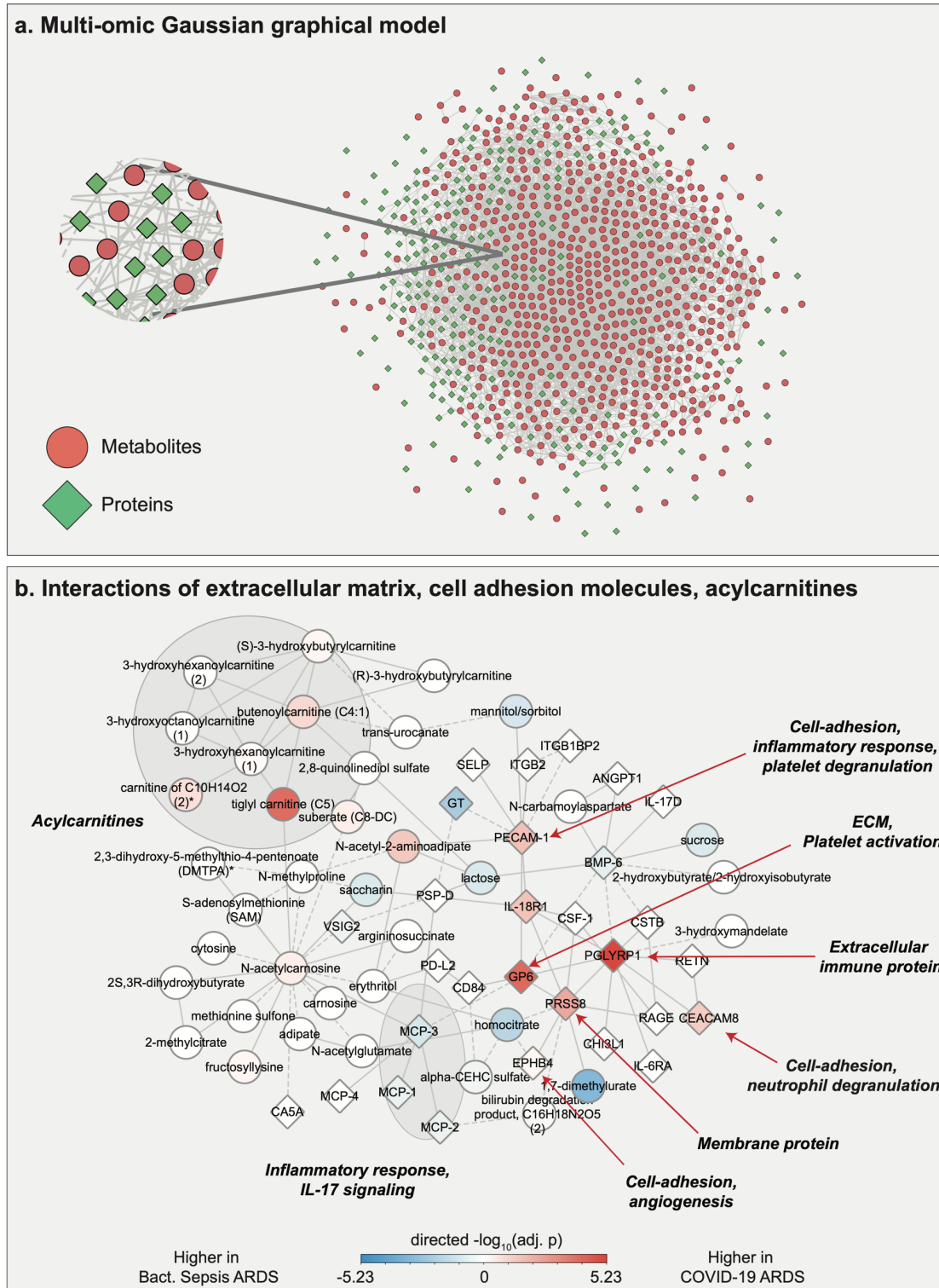
## 2.2. ARDS-related interaction of mitochondrial dysfunction and ECM organization

Predefined pathway annotations provide context for already well-characterized biological processes; however, the insights they provide into cross-omics associations are limited. Therefore, we generated a data-driven multi-omic interaction network based on Gaussian graphical models (GGM)<sup>32</sup>. In earlier studies, we have shown that partial correlation-based GGMs reconstruct valid biochemical interactions from omics data in an unbiased fashion and can even identify previously unknown interactions between molecules<sup>33–35</sup>. The data-driven network contained 3,566 statistically significant interactions between the 708 metabolites and 266 proteins (**Figure 2a**). It was then annotated using the molecules that were differentially abundant between ARDS groups. An interactive version of the network is available in **Supplementary File 1** for further exploration.

We then generated a subnetwork focusing on several processes that have previously been implicated in COVID-19, namely mitochondrial dysfunction<sup>36</sup>, coagulopathy via cell-adhesion molecules (CAMs) and platelet activation<sup>37,38</sup>. To this end, we chose two molecules belonging to these processes which were also among the top metabolomic and proteomic hits in Figure 1a: Tiglyl carnitine, an acylcarnitine that represents mitochondrial function<sup>39</sup>, and glycoprotein 6 (GP6), which is involved in the extracellular matrix (ECM) and the platelet activation pathway.

The subnetwork was constructed by including tiglyl carnitine, GP6, and all of their first- and second-degree network neighbors, i.e., nodes that were separated from the two molecules by one or two edges in the network. The resulting subnetwork consisted of 66 molecules (37 metabolites, 29 proteins) with 106 interactions among them (**Figure 2b**). Within this subnetwork, tiglyl carnitine and GP6 were connected via MCP-3 and N-acetylcarnosine. The neighborhood of tiglyl carnitine consisted of other acylcarnitines, including butenoylcarnitine (C4:1), (S)-3-hydroxybutyrylcarnitine, and 3-hydroxyhexanoylcarnitine, all of which were higher in COVID-19 compared to bacterial sepsis-induced ARDS. The neighborhood of GP6 consisted of additional proteins related to ECM or CAMs, including EPHB4, CECAM8, and PECAM1, all of which were higher in COVID-19 compared to bacterial sepsis-induced ARDS. The mediating inflammatory MCP-3 protein was connected to other MCP proteins, which were higher in bacterial sepsis-induced ARDS than in COVID-19.

Overall, within the subnetwork, we observed cross-omics connections between clusters of CAMs/ECM and a group of acylcarnitines, mediated by a group of inflammatory MCP proteins. We speculate that the underlying interplay of these depicted biological processes might play a role in ARDS pathogenesis.



**Figure 2: Multi-omic network and extracted ECM/CAM/acylcarnitine subnetwork.** a. Gaussian graphical model (GGM) of metabolites and proteins. Shapes and colors of the molecules in the network are based on the two omics types. b. Subnetwork extracted from the full multi-omic GGM, built around tiglyl carnitine and GP6. The observed molecular interactions suggest an interplay of ECM derangement, inflammation, and mitochondrial dysfunction in ARDS pathogenesis.

### 2.3. ARDS-specific heterogeneity of molecular associations across clinical manifestations

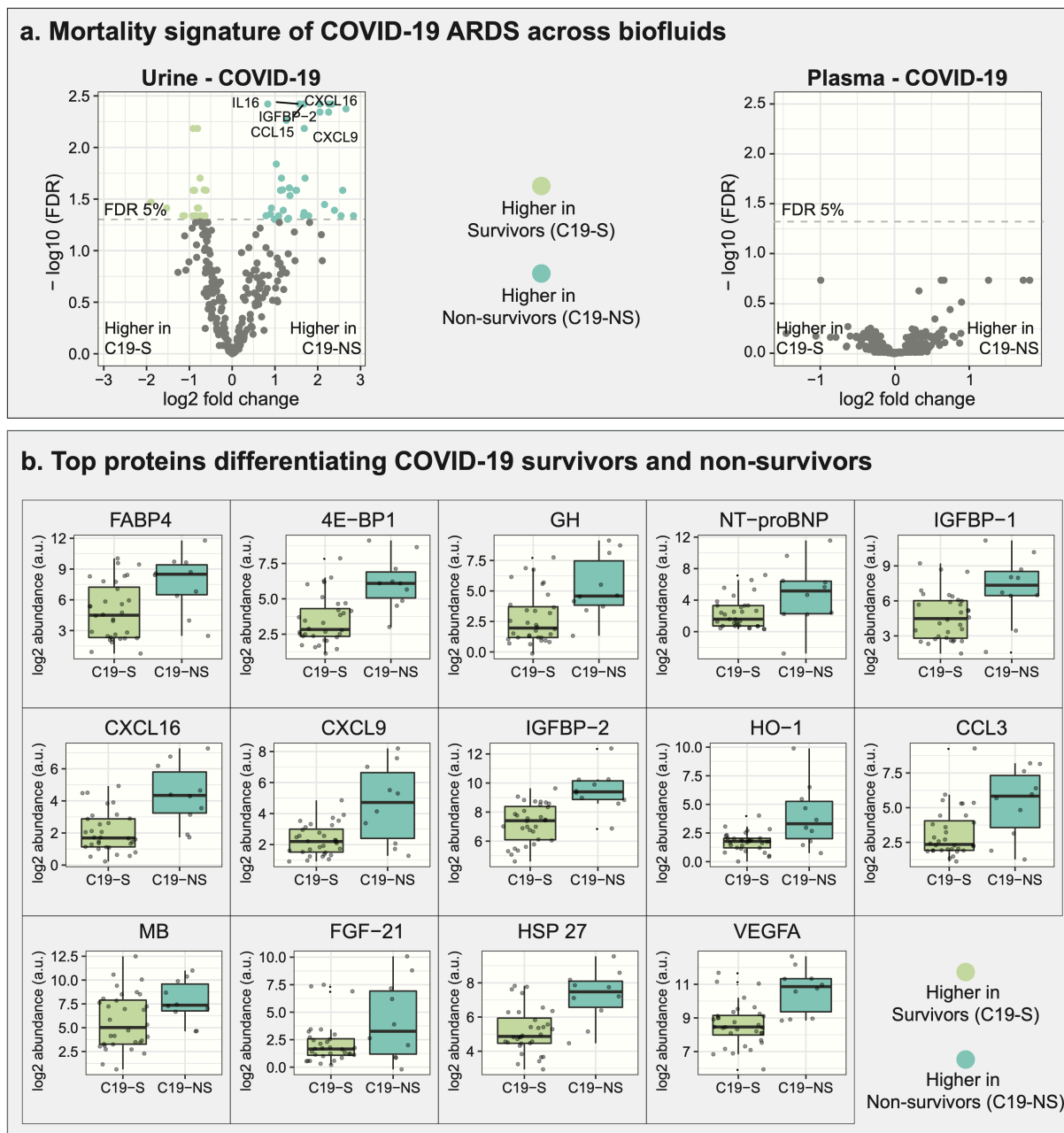
In the second part of our study, we tested ARDS group-specific molecular associations with four clinical manifestations: acute kidney injury (AKI), platelet count, patient's oxygen in arterial blood to the fraction of the oxygen in the inspired air (PaO<sub>2</sub>/FIO<sub>2</sub>) ratio, and mortality. In the bacterial sepsis-induced ARDS group, no significant associations with any of these clinical manifestations were identified (5% FDR). In COVID-19 ARDS, there were 10 molecules associated with AKI, including 8 metabolites and 2 proteins, no molecules associated with platelet count, 6 metabolites associated with PaO<sub>2</sub>/FIO<sub>2</sub>, and 61 molecules associated with mortality, including 1 metabolite and 60 proteins. Thus, a molecular comparison of heterogeneous presentations across the two ARDS groups was not feasible. Detailed results are available in **Supplementary Tables 3 and 4**.

In the following, we focused on the proteomic mortality signature distinguishing survivors and non-survivors of COVID-19. Among 60 proteins that were significantly different between survivors and non-survivors, 22 were higher in survivors, 38 higher in non-survivors (**Figure 3a, left**). Remarkably, in our recent plasma-based study<sup>15</sup>, we did not find any proteins that were associated with mortality in the same COVID-19 patients (**Figure 3a, right**).

For further investigation of the urine-based COVID-19 mortality signature, we selected significant proteins with log<sub>2</sub> fold changes larger than 2 (**Figure 3b**). Interestingly, several of these 14 proteins have previously been described as biomarkers of pathologies that are linked to ARDS. For example, NT-proBNP in the urine of preterm infants has been shown to inform about pulmonary hypertension<sup>40</sup>. IGFBP-2 is an indicator of pulmonary arterial hypertension (PAH)<sup>41</sup> and can predict a decline of kidney function in type 2 diabetes<sup>42</sup>. FABP4 has been implicated in proteinuria and has been discussed as a marker of kidney glomerular damage<sup>43</sup>. CXCL16 is considered a urinary marker of poor renal outcome in diabetic kidney disease<sup>44</sup>. HO-1 is a candidate biomarker for oxidative damage in obstructive nephropathy<sup>45</sup>. VEGFA leads to increased inflammation in severe COVID-19<sup>46</sup>.

Taken together, these mortality-associated proteins have been implicated in ARDS-linked manifestations, including kidney dysfunction, pulmonary hypertension, and inflammation<sup>8,47,48</sup>. This provides insights into the potential pathophysiological processes behind the development of severe ARDS.





**Figure 3. Proteomics-based mortality signature.** **a.** Differentially abundant proteins in COVID-19 survivors and non-survivors, as observed in two bodily fluids, urine and plasma. 60 proteins were significant in urine proteomic profiles, while none of the proteins measured in the plasma of the same patients were associated with mortality. **b.** Top 14 differential proteins from COVID-19 urine-based mortality signature with log<sub>2</sub> fold change larger than or equal to 2 at 5% FDR.

## 2.4. Conclusion

In this study, we presented a first urine-based multi-omic comparison of COVID-19 ARDS and non-COVID-19 ARDS. We compared 42 COVID-19 ARDS patients to 17 bacterial sepsis-induced ARDS patients using untargeted metabolomics (708 metabolites) and targeted proteomics (266 proteins). There were two main findings from our work. First, the multi-omic network approach highlighted the interplay of mitochondrial dysfunction and ECM derangement in ARDS pathogenesis. Second, we identified a proteomics-based mortality signature in COVID-19 ARDS patients. Notably, within the bacterial sepsis-induced ARDS group, no metabolites or proteins were found to be associated with any of the four clinical manifestations tested. In the following paragraphs, we discuss the two novel findings from our study.

Our multi-omic network-based analysis indicated an ARDS-related link between CAMs/ECM and mitochondrial dysfunction represented by acylcarnitines. In the analyzed subnetwork, the connections between these different biological processes were mediated by inflammatory proteins. Previous COVID-19 studies have already individually implicated these processes in ARDS, but have not proposed a link between these pathways in the context of ARDS<sup>26,36,39</sup>. Moreover, mechanistically, ECM, CAMs, and acylcarnitines have individually been linked with inflammation<sup>49-51</sup>. Our findings now highlight the potential synergy between these different cellular pathways in ARDS.

The proteomics-based mortality signature distinguishing COVID-19 survivors and non-survivors is another potentially novel finding from our study. Surprisingly, the mortality signal was absent in plasma proteomics profiles of the same patients. This could reflect frequent kidney involvement in severe COVID-19, which leads to the poor renal outcomes observed in our COVID-19 patients<sup>48</sup>. Moreover, the signature contains several proteins implicated in pathological processes that have been linked to ARDS, including inflammation, kidney dysfunction, and pulmonary hypertension<sup>8,47,48</sup>. In terms of clinical stratification approaches, a higher-powered study will be needed to assess whether machine learning models based on our signature are able to predict mortality in COVID-19 ARDS patients.

We recognize that our study design has several limitations. (1) Since the patients of the two ARDS groups were collected several years apart, differences in sample collection and storage protocols may lead to unaccountable variation across measurements. (2) Our cohort has a limited sample size ( $n = 59$ ), with imbalanced ARDS groups (42 COVID-19 versus 17 bacterial sepsis patients). This relatively small sample size could have led to false negatives in our analysis, especially within the bacterial sepsis group. (3) Since the coverage of the metabolomics and proteomics platforms is limited, there is potential for missed associations with unmeasured molecules. (4) Our study was limited to statistical associations in a single cohort since we did not have access to an independent cohort for replication.

In conclusion, we presented a first urine-based multi-omic analysis of COVID-19 ARDS compared to bacterial sepsis-induced ARDS. Our analysis shows molecular similarities and differences

between the two ARDS groups. The most striking finding was a proteomics-based mortality signature specifically for COVID-19 ARDS, which will require further investigation as a potential early biomarker for mortality.

### 3. Methods

#### 3.1. Patient Population

The cohort was derived from the Weill Cornell Biobank of Critical Illness (WC-BOCI) at WCMC/NYP. The process for recruitment, data collection, and sample processing have been described previously<sup>54–56</sup>. In brief, the recruits in the WC-BOCI database were patients admitted to the intensive care unit with valid consent between October 2014 to May 2021, including 59 patients with COVID-19 ARDS (n=42) and bacterial sepsis-induced ARDS (n=17). Clinical data such as demographics, vital signs, labs, and ventilator parameters were obtained through the Weill Cornell-Critical Care Database for Advanced Research (WC-CEDAR) and the Weill Cornell Medicine COVID Institutional Data Repository (COVID-IDR). Additional clinical data were obtained through manual abstraction from the electronic health records.

This cohort included 47 (79.7%) males and 12 (20.3%) females, with a median age of 58.3. The overall mortality rate was 27.1 %, with 10 out of 42 in COVID-19 ARDS and 6 out of 17 in bacterial sepsis-induced ARDS. 45.8% of patients suffered from acute kidney injury (AKI), with 15 out of 42 in COVID-19 ARDS and 12 out of 17 in bacterial sepsis-induced ARDS. The sequential organ failure assessment (SOFA) index was comparable between the two groups, with a median of 10 in the COVID-19 group and 9 in the bacterial sepsis group. Detailed demographics of the patient cohort are provided in **Supplementary Table 5**.

#### 3.2. Clinical manifestations

Below are the definitions used to diagnose the clinical manifestations used in this study.

*Acute Respiratory Distress Syndrome (ARDS)*. ARDS was assessed using the Berlin definition<sup>57</sup>, and followed by a review of the subject's history, arterial blood gas, and chest X-ray by two independent pulmonary and critical care attendings to adjudicate the diagnosis. For bacterial sepsis-induced ARDS, an additional criterion was used as outlined in The Third International Consensus Definitions for Sepsis and Septic Shock<sup>58</sup>. For diagnosis of COVID-19, a positive viral swab of the nasopharynx tested for SARS-CoV-2 via RT-PCR was required.

*Acute Kidney Injury (AKI)*. 'Kidney Disease: Improving Global Outcomes' definition (KDIGO) was used to diagnose AKI. To this end, either of the following criteria was required: (a) serum creatinine change of greater than or equal to 0.3 mg/dL within 48 hours, (b) serum creatinine greater than or equal to 1.5 times the baseline serum creatinine known or assumed to have occurred within the past 7 days, (c) urine output less than or equal to 0.5 mL/kg/hour for six hours<sup>59</sup>.

### **3.3. Sample handling**

Urine specimens were obtained from patients admitted to ICU at WCMC/NYP. Briefly, urine samples were centrifuged, and the supernatant was stored at -80°C until the omics profiling was performed. An electronic informed consent was obtained from all subjects for inclusion.

### **3.4. Proteomic profiling**

Proteomic profiling was performed by the Proteomics Core of Weill Cornell Medicine-Qatar using the Olink platform (Uppsala, Sweden)<sup>15</sup>. Briefly, manufacturer's instructions were followed to profile the samples using four panels including Inflammation, Cardiovascular II, and Cardiovascular III panels. Thorough quality assurance/quality control (QA/QC) was performed to monitor the assay's incubation, extension, and detection steps. For (Ct) value extraction, Fluidigm's reverse transcription-polymerase chain reaction (RT-PCR) analysis software was used at a quality threshold of 0.5 and linear baseline correction. Further processing of Ct values was performed using the Olink NPX manager software (Olink, Uppsala, Sweden).

### **3.5. Metabolomic profiling**

Metabolic profiling was performed by Metabolon, Inc (Morrisville, NC) using ultrahigh performance liquid chromatograph-tandem mass spectroscopy (UPLC-MS/MS)<sup>15</sup>. Briefly, samples were subjected to methanol extraction and then divided into four aliquots for each of the mass spectroscopic methods. Rigorous quality assurance/quality control (QA/QC) was performed to monitor instrument performance and aid in chromatographic alignment. The four mass spectroscopic methods used were optimized for acidic positive ion hydrophilic compounds, acidic positive ion hydrophobic compounds, and basic negative ions, the fourth aliquot was analyzed via negative ionization. For metabolite identification, Metabolon's proprietary software was used to deliver high-quality abundances of metabolites.

### **3.6. Data processing**

Metabolomic and proteomic profiles were preprocessed before downstream analysis: Molecules with more than 25% missing values were removed, leaving 708 out of 1,112 metabolites and 266 out of 276 proteins. Probabilistic quotient normalization<sup>60</sup> was used to correct sample-wise variation in the data. Data was log<sub>2</sub> transformed, followed by k-nearest-neighbor-based imputation<sup>61</sup> for the remaining missing values. Abundance levels of the following ten proteins were measured in duplicates by Olink panels and were therefore averaged: CCL3, CXCL1, FGF-21, FGF-23, IL-18, IL-6, MCP-1, OPG, SCF, and uPA. All data processing was performed using the maplet R package<sup>62</sup>.

### **3.7. Differential analysis of molecules**

For association analysis, we used linear models with metabolites/proteins as the dependent variable and diagnosis/clinical manifestations as independent variables. Further factors such as

age, sex, and BMI were not used as covariates in the models, since they are considered determinants of disease severity themselves<sup>63</sup>. To control the false discovery rate, the Benjamini-Hochberg (BH) method<sup>64</sup> was used to correct p-values. All analyses were performed using the maplet R package<sup>62</sup>.

### 3.8. Pathway annotation and filtering

For functional annotation of the differently abundant molecules, we used Metabolon's 'sub-pathway' groups and signaling pathways from KEGG<sup>25</sup> for metabolites and proteins, respectively. **Supplementary Table 2** contains the complete list of annotations. For our analysis, we considered Metabolon's sub-pathways with the term 'metabolism' and non-disease KEGG pathways with at least 3 significant molecules.

### 3.9. Multi-Omic network inference

To generate a multi-omic data-driven network we created a Gaussian graphical model (GGM) using the GeneNet R package<sup>32</sup>. GGMs are a partial correlation-based approach for identifying statistical connections among the molecules. To construct the network, pair of molecules (nodes) with significant partial correlations at 5% FDR were included and were connected with an edge. Following this, these nodes were annotated based on the statistical association results between the ARDS groups. To this end, a  $p_{score}$  was computed using the following formula:  $p_{score} = -\log_{10}(p.adj) \cdot d$ , where  $p.adj$  is the adjusted p-value of the association, and  $d$  is the direction (-1/1) of the association based on test statistic (positive or negative association with the outcome). This score was used to color the nodes in the network.

## References

1. UNDP. Everyone Included: Social Impact of COVID-19 | DISD. *United Nations* <https://www.un.org/development/desa/dspd/everyone-included-covid-19.html>.
2. Kimberly Chriscaden. Impact of COVID-19 on people's livelihoods, their health and our food systems. *World Health Organization* 4–7 <https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people's-livelihoods-their-health-and-our-food-systems> (2020).
3. S, H. & RK, M. The acute respiratory distress syndrome: from mechanism to translation. *J. Immunol.* **194**, 855–860 (2015).
4. Grant, R. A. *et al.* Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nat.* 2021 5907847 **590**, 635–641 (2021).
5. Helms, J. *et al.* High risk of thrombosis in patients with severe SARS-CoV-2 infection: a multicenter prospective cohort study. *Intensive Care Med.* **46**, 1089–1098 (2020).
6. Hue, S. *et al.* Uncontrolled Innate and Impaired Adaptive Immune Responses in Patients with COVID-19 Acute Respiratory Distress Syndrome. *Am. J. Respir. Crit. Care Med.* **202**, 1509–1519 (2020).
7. Bain, R. *et al.* Clinical characteristics of SARS-CoV-2 infection in children with cystic fibrosis: An international observational study. *J. Cyst. Fibros.* **20**, 25–30 (2021).
8. Brault, C. *et al.* COVID-19- versus non-COVID-19-related acute respiratory distress syndrome: Differences and similarities. *Am. J. Respir. Crit. Care Med.* **202**, 1301–1304 (2020).
9. Dostálová, V. & Dostál, P. Acute respiratory distress syndrome. *Vnitr. Lek.* **65**, 193–203 (2019).
10. Robinson, M. J. & Krasnodembskaya, A. D. Therapeutic targeting of metabolic alterations in acute respiratory distress syndrome. *Eur. Respir. Rev.* **29**, 1–10 (2020).
11. Levitt, J. E. & Rogers, A. J. Proteomic study of acute respiratory distress syndrome: current knowledge and implications for drug development. *Expert Rev. Proteomics* **13**, 457–469 (2016).
12. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst.* **12**, 23-40.e7 (2021).
13. Shen, B. *et al.* Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59-72.e15 (2020).
14. van de Veerdonk, F. L. *et al.* A guide to immunotherapy for COVID-19. *Nat. Med.* **28**, 39–50 (2022).
15. Batra, R. *et al.* Multi-omic comparative analysis of COVID-19 and bacterial sepsis-induced ARDS. *medRxiv* 2022.05.16.22274587 (2022) doi:10.1101/2022.05.16.22274587.
16. Wu, J. & Gao, Y. Physiological conditions can be reflected in human urine proteome and metabolome. *Expert Rev. Proteomics* **12**, 623–636 (2015).
17. Berry, M. *et al.* Proteomics analysis reveals IGFBP2 as a candidate diagnostic biomarker for heart failure. *IJC Metab. Endocr.* **6**, 5–12 (2015).

18. Aregger, F. *et al.* Identification of IGFBP-7 by urinary proteomics as a novel prognostic marker in early acute kidney injury. *Kidney Int.* **85**, 909–919 (2014).
19. Gisewhite, S., Stewart, I. J., Beilman, G. & Luszczek, E. Urinary metabolites predict mortality or need for renal replacement therapy after combat injury. *Crit. Care* **25**, 1–14 (2021).
20. Currie, G. E. *et al.* Urinary proteomics for prediction of mortality in patients with type 2 diabetes and microalbuminuria. *Cardiovasc. Diabetol.* **17**, 1–8 (2018).
21. Li, Y. *et al.* Urine proteome of COVID-19 patients. *Urine* **2**, 1–8 (2020).
22. Bi, X. *et al.* Proteomic and metabolomic profiling of urine uncovers immune responses in patients with COVID-19. *Cell Rep.* **38**, 110271 (2022).
23. Li, Y. *et al.* Multi-platform omics analysis reveals molecular signature for COVID-19 pathogenesis, prognosis and drug target discovery. *Signal Transduct. Target. Ther.* **6**, 1–11 (2021).
24. Tian, W. *et al.* Immune suppression in the early stage of COVID-19 disease. *Nat. Commun.* **11**, 5859 (2020).
25. Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**, D109–14 (2012).
26. Li, Y. *et al.* Urine Proteome of COVID-19 Patients. *medRxiv* (2020) doi:10.1101/2020.05.02.20088666.
27. Hou, X. *et al.* Serum Protein Profiling Reveals a Landscape of Inflammation and Immune Signaling in Early-stage COVID-19 Infection. *Mol. Cell. Proteomics* **19**, 1749–1759 (2020).
28. Grimes, J. M. & Grimes, K. V. p38 MAPK inhibition: A promising therapeutic approach for COVID-19. *J. Mol. Cell. Cardiol.* **144**, 63–65 (2020).
29. Kalil, A. C. *et al.* Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. *N. Engl. J. Med.* **384**, 795–807 (2021).
30. Guimarães, P. O. *et al.* Tofacitinib in Patients Hospitalized with Covid-19 Pneumonia. *N. Engl. J. Med.* **385**, 406–415 (2021).
31. Montaldo, C. *et al.* Multi-omics approach to COVID-19: a domain-based literature review. *J. Transl. Med.* **19**, 1–18 (2021).
32. Schäfer, J. & Strimmer, K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* **4**, 1–30 (2005).
33. Krumsiek, J., Suhre, K., Illig, T., Adamski, J. & Theis, F. J. Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Syst Biol* **5**, 21 (2011).
34. Do, K. T. *et al.* Phenotype-driven identification of modules in a hierarchical map of multifluid metabolic correlations. *npj Syst. Biol. Appl.* **3**, 1–12 (2017).
35. Benedetti, E. *et al.* Network inference from glycoproteomics data reveals new reactions in the IgG glycosylation pathway. *Nat. Commun.* **2017 81 8**, 1–15 (2017).
36. Ajaz, S. *et al.* Mitochondrial metabolic manipulation by SARS-CoV-2 in peripheral blood



- mononuclear cells of patients with COVID-19. *Am. J. Physiol. - Cell Physiol.* **320**, C57–C65 (2021).
37. M, T. *et al.* Elevated Expression of Serum Endothelial Cell Adhesion Molecules in COVID-19 Patients. *J. Infect. Dis.* **222**, 894–898 (2020).
  38. Sriram, K. & Insel, P. A. Inflammation and thrombosis in covid-19 pathophysiology: Proteinase-activated and purinergic receptors as drivers and candidate therapeutic targets. *Physiol. Rev.* **101**, 545–567 (2021).
  39. Knottnerus, S. J. G. *et al.* Disorders of mitochondrial long-chain fatty acid oxidation and the carnitine shuttle. *Rev. Endocr. Metab. Disord.* **19**, 93–106 (2018).
  40. Naeem, B. *et al.* Urinary NT-proBNP as a potential noninvasive biomarker for screening of pulmonary hypertension in preterm infants: a pilot study. *J. Perinatol.* **40**, 628–632 (2020).
  41. Yang, J. *et al.* Insulin-like growth factor binding protein-2: A new circulating indicator of pulmonary arterial hypertension severity and survival. *BMC Med.* **18**, 1–11 (2020).
  42. Narayanan, R. P. *et al.* IGFBP2 is a biomarker for predicting longitudinal deterioration in renal function in type 2 diabetes. *Endocr. Connect.* **1**, 95–102 (2012).
  43. Okazaki, Y. *et al.* Urinary excretion of fatty acid-binding protein 4 is associated with albuminuria and renal dysfunction. *PLoS One* **9**, (2014).
  44. Lee, Y. H. *et al.* Urinary chemokine C-X-C motif ligand 16 and endostatin as predictors of tubulointerstitial fibrosis in patients with advanced diabetic kidney disease. *Nephrol. Dial. Transplant* **36**, 295–305 (2021).
  45. Li, Z. *et al.* Urinary heme oxygenase-1 in children with congenital hydronephrosis due to ureteropelvic junction obstruction. *Biomarkers* **17**, 471–476 (2012).
  46. C, H. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet (London, England)* **395**, 497–506 (2020).
  47. Revercomb, L., Hanmandlu, A., Wareing, N., Akkanti, B. & Karmouty-Quintana, H. Mechanisms of Pulmonary Hypertension in Acute Respiratory Distress Syndrome (ARDS). *Front. Mol. Biosci.* **7**, 624093 (2020).
  48. Legrand, M. *et al.* Pathophysiology of COVID-19-associated acute kidney injury. *Nat. Rev. Nephrol.* **2021 1711** **17**, 751–764 (2021).
  49. Sorokin, L. The impact of the extracellular matrix on inflammation. *Nat. Rev. Immunol.* **2010 1010** **10**, 712–723 (2010).
  50. González-Amaro, R., Díaz-González, F. & Sánchez-Madrid, F. Adhesion molecules in inflammatory diseases. *Drugs* **56**, 977–988 (1998).
  51. Rutkowsky, J. M. *et al.* Acylcarnitines activate proinflammatory signaling pathways. *Am. J. Physiol. - Endocrinol. Metab.* **306**, E1378 (2014).
  52. Price, D. R. *et al.* Angiotensin 2 Is Associated with Vascular Necroptosis Induction in Coronavirus Disease 2019 Acute Respiratory Distress Syndrome. (2022) doi:10.1016/j.ajpath.2022.04.002.
  53. Ricciotti, E. & Fitzgerald, G. A. Prostaglandins and Inflammation. *Arterioscler. Thromb. Vasc. Biol.* **31**, 986 (2011).

54. Finkelsztein, E. J. *et al.* Comparison of qSOFA and SIRS for predicting adverse outcomes of patients with suspicion of sepsis outside the intensive care unit. *Crit. Care* **21**, (2017).
55. Dolinay, T. *et al.* Inflammasome-regulated cytokines are critical mediators of acute lung injury. *Am. J. Respir. Crit. Care Med.* **185**, 1225–1234 (2012).
56. Schenck, E. J. *et al.* Circulating cell death biomarker TRAIL is associated with increased organ dysfunction in sepsis. *JCI insight* **4**, (2019).
57. Ranieri, V. M. *et al.* Acute respiratory distress syndrome: The Berlin definition. *JAMA - J. Am. Med. Assoc.* **307**, 2526–2533 (2012).
58. Singer, M. *et al.* The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* **315**, 801–810 (2016).
59. Khwaja, A. KDIGO clinical practice guidelines for acute kidney injury. *Nephron. Clin. Pract.* **120**, (2012).
60. Dieterle, F., Ross, A., Schlotterbeck, G. & Senn, H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabolomics. *Anal. Chem.* **78**, 4281–4290 (2006).
61. Do, K. T. *et al.* Characterization of missing values in untargeted MS-based metabolomics data and evaluation of missing data handling strategies. *Metabolomics* **14**, (2018).
62. Chetnik, K. *et al.* maplet: an extensible R toolbox for modular and reproducible metabolomics pipelines. *Bioinformatics* **38**, 1168–1170 (2022).
63. Docherty, A. B. *et al.* Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* **369**, (2020).
64. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

## Study Approval

The study was approved by the institutional review board at Weill Cornell Medicine (protocol number 22-03024534). Written informed consent was received before participation by all patients, except when the institutional review board approved a waiver of informed consent (e.g., for the use of discarded samples and de-identified patient data).

## Data and code availability

The preprocessed data used in this study can be downloaded at <https://doi.org/10.6084/m9.figshare.20260998.v1>

All R scripts to generate the tables and figures of this paper are available at <https://github.com/krumnsieklab/covid-ards-urine>

## Funding

JK and RB are supported by the National Institute of Aging of the National Institutes of Health under awards 1U19AG063744 and R01AG069901-01. ES is supported by NHLBI K23 HL151876. OA is supported by the NIH NIDDK K08 DK114558. FS was strongly supported by the Biomedical Research Program at Weill Cornell Medicine in Qatar, a program funded by the Qatar Foundation.

## Conflict of Interest

A.M.K.C. is a cofounder and equity stockholder for Proterris, which develops therapeutic uses for carbon monoxide. A.M.K.C. has a use patent on CO. Additionally, A.M.K.C. has a patent in COPD. ES consults for Axle informatics regarding COVID vaccine clinical trials through NIAID. JK holds equity in Chymia LLC and IP in PsyProtix and is cofounder of iollo.

## Author contributions

RB, RU, AMKC, OMA, MEC, and JK designed the study. SJC and ES contributed to study design discussions. KS and FS contributed to proteomics profiling. SAM, LGGE, RU, WS, and KLH extracted de-identified clinical data. LGGE and EP collected samples. EB, MB, and KC contributed to statistical discussions. RB performed the analysis. RB, RU, OMA, MEC, and JK interpreted the results. RB, RU, MEC, and JK wrote the manuscript. MEC and JK supervised the study. All authors approved the final manuscript.