Distribution agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restriction as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of this thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Ralitza Martin MD Date

Metabolomic Analysis in Sepsis Induced Acute Respiratory Distress Syndrome

By

Ralitza Martin

Master of Science

Clinical Research

_____(Advisor's signature)

Gregory Martin, MD MS

Advisor

_____(Advisors's signature)

Annette Esper, MD MS

Advisor

_____(Member's signature)

Amita Manatunga, PhD

Committee Member

_____(Member's signature)

Beau Bruce, MD PhD

Committee Member

Accepted

_____

Lisa A. Tedesco, PH.D.

Dean of the James T. Laney School of Graduate Studies

_____

Date

Metabolomic Analysis in Sepsis Induced Acute Respiratory Distress Syndrome

By

Ralitza Martin

MD, Universitaet des Saarlandes, Germany 2006

Advisor: Annette Esper, MD, MS

Advisor: Gregory Martin, MD, MS

An abstract of

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Clinical Research

2014

Abstract

Metabolomic Analysis in Sepsis Induced Acute Respiratory Distress Syndrome

By Ralitza Martin

Rationale: Metabolomic analysis is a novel approach in critical illness providing rapid analysis of metabolites from various biological materials, such as plasma and bronchoalveolar lavage fluid (BALF), with minimal sample preparation. This study tested if metabolomics could discriminate sepsis patients according to development of ARDS and mortality.

Methods: Prospective cohort of adult ICU patients at Grady Memorial Hospital meeting the ACCP/SCCM definition of severe sepsis or septic shock. Following informed consent, each subject underwent blood sampling and bronchoscopy with BAL. BALF and plasma were analyzed on a LCMS metabolomics platform using C18 chromatography and high-resolution mass spectrometry and analyzed using LIMMA to identify differentially expressed metabolites (DEMs) after adjusting for false discovery rate (FDR). Classification accuracies were evaluated using 10-fold cross-validation (CV) using support vector machine.

Results: Metabolomic analysis was performed on 79 BALF and plasma samples, 42 had ARDS. The mean age of the cohort was 53.9 years, 59 % male, 68% black. The overall mortality for the cohort was 48%. There was no mortality difference between ARDS and non-ARDS patients (42.9% vs. 54.1%, p=0.32). BALF FDR analysis between survivors and non-survivors of sepsis revealed 37 DEMs with 85.9% 10-fold CV classification accuracy. Top 500 metabolites from the network analysis were found to be significantly enriched in the lipid and drug metabolism, and aminoacid biosynthesis pathways. BALF analysis between ARDS and non-ARDS patients showed 35 DEMs with 92.1% CV accuracy, with pathways for aminoacid and drug metabolism, significantly enriched. Analysis of plasma did not reveal any significant DEMs in regards to development of ARDS or mortality after FDR correction.

Conclusion: While BALF metabolites were able to discriminate between sepsis non-survivors and survivors, and those developing ARDS or not, plasma metabolites could not. Sampling of the lung compartment may be necessary to understand the metabolic changes occurring in ARDS that may be linked to survival.

Metabolomic Analysis in Sepsis Indiced Acute Respiratory Distress Syndrome

By

Ralitza Martin

MD, Universitaet des Saarlandes, Germany 2006

Advisor: Annette Esper, MD, MSc

Advisor: Gregory Martin, MD, MSc

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Clinical Research
2014

TABLE OF CONTENTS

INTRODUCTION

The terms Acute Lung injury (ALI) and Acute Respiratory Distress Syndrome (ARDS) refer to an acute abnormality of both lungs characterized by bilateral inflammatory pulmonary infiltrates and impaired oxygenation. Military physicians have used various terms, such as “wet lung” and “shock lung”, to describe the edematous lung injury caused by trauma that they witnessed. However, in 1967 Ashbaugh et al. coined the term “respiratory distress syndrome” to describe the constellation of acute onset of tachypnea, hypoxemia, diffuse pulmonary infiltrates and loss of lung compliance resulting in high short term mortality. [1, 2]

The terms ALI and ARDS achieved a consensus definition during the American-European Consensus Conference (AECC) in 1994, which allowed for a streamlined research effort into the epidemiology, pathophysiology and treatment of ALI/ARDS. This definition was later revised in 2012 at a conference in Berlin to provide a more feasible, reliable and prognostic criteria (Table 1).[3] However, currently most of the existing evidence is based on the prior AECC definition.

There are approximately 200,000 cases of ARDS in the US each year, accounting for over 2 million ICU days and 3.6 million hospital days annually and the incidence in the US maybe slightly higher than other countries [8, 9]. The mortality of ARDS remains high at 25-40%.[9, 10] Because of the high mortality and substantial variability in outcomes, identification of risk factors for mortality is important to determine prognosis and guide clinical decisionmaking.[2] Various biomarkers have been studied in ARDS (e.g., surfactant proteins, cytokines, and

markers of pulmonary epithelial and endothelial injury), but the results have been rather disappointing.[11] Establishing a relationship between measurable biological processes and clinical outcomes is crucial to advancing clinical trials in ARDS and expanding our potential treatments for this complex syndrome and we need to expand our efforts to include novel approaches such as metabolomics.[12]

BACKGROUND

ARDS is a consequence of an alveolar injury leading to diffuse alveolar damage. The injury, triggered by various predisposing conditions (Table 2), causes release of pro-inflammatory cytokines which recruit neutrophils to the lung where they get activated and release toxic mediators that damage the capillary endothelium and alveolar epithelium causing it to become "leaky". This leads to fluid, protein, and debris from degenerating cells pouring into the alveoli. [4] In addition, the surfactant, which keeps the alveoli open, is lost, leading to alveolar collapse. This lung injury results in impaired gas exchange, decreased lung compliance and increased pulmonary arterial pressure. Patients with ARDS tend to progress through three discrete pathologic stages: exudative, proliferative, and fibrotic. [5] Although many predisposing factors have been identified (Table 2), **sepsis is the most common cause for ARDS.** [6] Other risk factors include history of alcohol abuse, obesity and admission severity of illness. [7]

Sepsis is a clinical syndrome that complicates a severe infection and is characterized by signs of systemic inflammation occurring in tissues remote from the infection. Systemic inflammatory response syndrome (SIRS), sepsis, severe sepsis, and septic shock are a continuum of this clinical syndrome and were initially defined in 1991 by a consensus panel convened by the American College of Chest Physicians (ACCP) and Society of Critical Care Medicine (SCCM) (ACCP/SCCM guidelines 1992). These were later modified to provide exact definitions for sepsis vs. septic shock. [13]

Risk factors for sepsis include bacteremia (presence of bacteria in the blood), advanced age (>65 years), immunocompromised state (e.g., medicationinduced immune compromise, malignancy, renal failure, AIDS).

The **incidence** of sepsis is estimated to be ~80 cases per 100,000 and has risen over the past decades - thought to be mainly due to advancing age, immunosuppression, multi-drug resistant infection. [14, 15]

As mentioned, sepsis is a clinical syndrome and provides a spectrum and we'll briefly describe the definitions in this spectrum.

SIRS -> sepsis -> severe sepsis -> septic shock -> multiorgan dysfunction syndrome (MODS)

SIRS is a clinical syndrome that is a form of dysregulated inflammation and can be associated with both infectious processes and non-infectious insults such as autoimmune disorders, pancreatitis, burns, and surgery. Sometimes SIRS cannot be distinguished from early sepsis so a search for a source of infection should be initiated when SIRS symptoms are present.[16]

Sepsis is a clinical syndrome that is characterized by presence of one or more SIRS symptoms and a documented or suspected infection (Dellinger).[16] Symptoms of SIRS include:

General variables

- Temperature >38.3 or <36°C
- Heart rate >90 beats/min

- Tachypnea, respiratory rate >20 breaths per minute
- Altered mental status
- Hyperglycemia (plasma glucose >140 mg/dL) in the absence of diabetes

Inflammatory variables

- Leukocytosis (white blood cell count (WBC) count of >12,000/μL) or leucopenia (WBC <4,000/μL)
- Plasma procalcitonin more than two standard deviations above the normal value

Hemodynamic variables

- Arterial hypoxemia (arterial oxygen tension (PaO2)/fraction of inspired oxygen(FiO2) <300

Organ dysfunction variables

- Acute oliguria (urine output <0.5 mL/kg/hr) for at least two hours despite adequate fluid resuscitation)
- Coagulation abnormalities (international normalized ratio (INR) >1.5 or activated partial thromboplastin time (aPTT)>60 seconds)
- Thrombocytopenia (platelet count <100,000/μL)
- Hyperbilirubinemia (plasma total bilirubin >4 mg/dL)

Tissue perfusion variables

- Elevated lactate (>1 mmol/L)

Severe sepsis refers to sepsis induced tissue hypoperfusion or organ dysfunction thought to be secondary to the infection and includes any of the following variables:

- Sepsis induced hypotension (systolic blood pressure <90 mmHg or mean arterial pressure <70 mmHg in the absence of other causes for hypotension) ○ Lactate above upper limit of laboratory normal ○ Urine output <0.5 mL/kg/h) for at least two hours despite adequate fluid resuscitation ○ Acute Respiratory Distress Syndrome (ARDS) with PaO2/FiO2 ratio <250 in the absence of pneumonia as the infection source ○ Creatinine >2mg/dL ○ Bilirubin >4 mg/dL ○ Platelet count <100,000/μL ○ Coagulopathy (INR>1.5)

Septic shock is a defined as sepsis-induced hypotension (low blood pressure) despite adequate fluid resuscitation which is defined as infusion of 30 ml/kg of crystalloid solution.

Multiple organ dysfunction syndrome (MODS) refers to progressive organ dysfunction in an acutely ill patient.

The lung is the organ most commonly involved in severe sepsis and MODS [17]. Despite the high incidence and poor prognosis of ARDS, it remains challenging to identify patients who are at higher risk and to prognosticate once the diagnosis is made.[9, 18] Progress in these areas can have a great impact by facilitating

studies aimed at prevention, identifying patients more accurately so they can be enrolled in clinical trials, and guiding therapeutic choices. Investigation of biological markers by sampling the pulmonary compartment and peripheral blood has helped us advance our knowledge of the pathogenesis of ARDS and provided more information on risk factors, diagnosis and prognosis.

A **biomarker** can be anything that is objectively measured, such as vital sign, laboratory value, or finding on an imaging study. Biomarkers are used to describe both normal and pathological processes and can be linked to clinical outcomes.[19, 20] In order to hasten the development of effective therapy for Acute Respiratory Distress Syndrome (ARDS), the National Heart, Lung, and Blood Institute, National Institutes of Health, initiated a clinical network to carry out multi-center clinical trials of ARDS treatments. The ARDS Network clinical trials group has provided large sample sizes of well phenotyped patients along with serial blood and in some cases bronchoalveloar lavage fluid (BALF).

In addition to potential treatments many potential biomarkers have been studied:

- markers of injury and activation of the **pulmonary endothelium** such as angiopoietin-2, selectins, von Willebrand Factor antigen, intercellular adhesion molecule-1, vascular endothelial growth factor (VEGF)
- markers of injury of the **pulmonary epithelium** such as surfactant proteins, receptor for advanced glycation end-products (RAGE), Clara cell protein
- Markers of **inflammation and coagulation**- cytokines, high-mobility group box 1 protein, plasminogen activator inhibotro-1, protein C.

Unfortunately, these studied biomarkers do not possess either great sensitivity, specificity or ease of measurement [21]. The development of these biomarkers has been mostly hypothesis driven and focused on studying our current understanding of the pathogenesis of ARDS. This targeted approach can potentially miss potential mediators of lung injury that are yet to be recognized as important. Objective selection methodologies, including metabolomics, proteomics, gene expression analysis, and genome-wide association studies can play an important role to identify potential new biomarkers of ARDS, as well as new mediators and pathways.

Metabolomics is the study of the dynamics of small molecules (metabolites) in living systems. Metabolomics is a rapidly expanding new field of –omics. The key concept is that changes in the proteome, transcriptome or genome are reflected in the metabolome as alterations in metabolite concentrations in biological fluids or tissues and measuring these metabolites can generate a

"snapshot" of the chemical and molecular composition in selected compartments of the body.

The theory is that things that influence the system (e.g., drugs, diseases, and environmental causes) lead to change in the metabolic profile. By capturing these alterations in metabolic concentrations, we can associate them to certain disease states, use them to develop biomarkers that could lead to early diagnosis, delineate disease phenotypes and also give us a better understanding of the pathophysiology. Quantitative metabolite data can be used to identify the related genes and the potential advantage is that the approach is not based on pre-conceived ideas and is hypothesis generating.

There is a range of metabolite-detecting technologies.

Separation methods:

- Gas chromatography (GC), especially when interfaced with mass spectrometry (GC-MS), is one of the most widely used and powerful methods. It offers very high chromatographic resolution, but requires chemical derivatization for many biomolecules: only volatile chemicals can be analyzed without derivatization. Some large and polar metabolites cannot be analyzed by GC.

Detection methods:

- Mass spectrometry is used to identify and to quantify metabolites after separation by GC or liquid chromatography (LC). GC-MS is the most 'natural' combination, and was the first to be developed. In addition, mass spectral fingerprint libraries exist or can be developed that allow identification of a metabolite according to its fragmentation pattern. MS is both sensitive and can be very specific. There are also a number of studies that use MS as a stand-alone technology: the sample is infused directly into the mass spectrometer with no prior separation, and the MS serves to both separate and to detect metabolites.
- Nuclear magnetic resonance (NMR) spectroscopy- the only detection technique that does not rely on separation of the analytes, and the sample can thus be recovered for further analyses. All kinds of small molecule metabolites can be measured simultaneously - in this sense, NMR is close to being a universal detector. The main advantages of NMR are high analytical

reproducibility and simplicity of sample preparation. Practically, however, it is relatively insensitive compared to mass spectrometry-based techniques. [22, 23]

To our knowledge the use of metabolomics in ARDS is scarce, emphasizing the need for further investigation in this patient population. In fact, there are a just a few studies that have investigated the use of metabolomics in ARDS as a tool to diagnose and prognosticate. In a small pilot study Frank et al. studied exhaled breath condensate of 6 patients and 2 controls and found that the metabolic profiles of ARDS patients were significantly different compared to at risk controls.[24]

Using NMR Stringer et al. were able to elucidate different metabolic patterns in the plasma of 13 patients with mild ARDS compared to healthy subjects.[25] They identified total glutathione, adenosine, phosphatidylserine, and sphingomyelin as the main differentiating metabolites. However, it was unclear if these were specific for ARDS or just markers of inflammation since all patients had sepsis and ARDS and they were compared to healthy, non-smoking volunteers.

Izquierdo-Garcia et al. studied 28 patients with the diagnosis of pneumonia by H1N1 influenza virus (12 of them diagnosed of ARDS). Plasma samples were analyzed by high-resolution magic angle spinning 1H-NMR spectroscopy.[26] Unsupervised principal components analysis (PCA) and supervised partial least squares (PLS) analysis were performed on the processed spectra to highlight

biomarkers of ARDS. NMR-based metabolic profiling discriminated characteristics between H1N1 influenza patients with ARDS from those without
ARDS. A PLS predictive model classified ARDS samples with a success rate of 91% (sensitivity 70%, specificity 93%).

Esper et al. investigated plasma of a total of 120 sepsis patients (60 with mild ARDS vs. 60 without ARDS) and showed different metabolic profiles where the top 1% of features accounting for 95% of the separation included arginine, arachidonate metabolism, bile acid metabolism, and sulfur amino acid metabolism.[27]

These studies indicate that metabolomics can be a useful tool in ARDS for diagnostic and prognostic purposes. Both targeted and global profiling has been successfully carried out in mass spectrometry based metabolomics [28].

Given this prior information we hypothesized that in patients with severe sepsis there are differences in metabolic profiles between survivors and non-survivors and those who do and do not develop ARDS, and we aimed to identify candidate metabolites that differentiate survivors versus non-survivors and ARDS versus non-ARDS patients and also help us better understand the pathophysiology of ARDS.

METHODS

Study design

Exploratory prospective cohort study

Study population

The study obtained approval from the Emory University institutional review board (IRB). All mechanically ventilated patients in the adult Intensive Care Units (ICU) at Grady Memorial Hospital, Atlanta, GA, USA were screened for severe sepsis or septic shock. The ICUs were screened daily for patients that meet the inclusion criteria. Informed consent was obtained from patients' next of kin or healthcare power of attorney.

Inclusion criteria: Mechanically ventilated patients were eligible for enrollment within the first 72 hours of severe sepsis or septic shock development. Patients were required to meet all the inclusion criteria within a 24-hour period. The definition of severe sepsis and septic shock is based on the American College of Chest Physicians/Society of Critical Care Medicine criteria for sepsis [16]:

a. Patients must meet at least two of the following criteria for SIRS: 1) temperature > 38C; 2) white blood cell count >$12X10^9$/L or <$4X10^9$, or presence of >10% bands/immature neutrophils; 3) respiratory rate >20 breaths/min; 4) heart rate >90 beats/min

b. Patients had a clearly defined source of infection
c. Patients must meet the criteria for severe sepsis: sepsis associated with organ dysfunction, hypoperfusion, or hypotension: hypoperfusion and perfusion abnormalities may include, but are not limited to, lactic acidosis, oliguria, or an acute alteration in mental status
d. Septic shock is defined as meeting at least one of the following: a) systolic blood pressure <90 mm Hg for >1 hr; b) the requirement for vasopressor therapy, excluding dopamine at a dose <5 μg/kg/min.

Exclusion Criteria: Patients were not enrolled if no informed consent could be obtained, if greater than 72 hours have lapsed since sepsis criteria were met or if patients were minors (age<18).

Primary exposures:

Severe sepsis/ septic shock and mechanical ventilation.

Primary outcomes:

The primary outcomes for the study were development of ARDS and 28-day mortality. Patients were considered to have ARDS if they met the Berlin definition criteria: onset within 7 days of an inciting event (in our case sepsis), bilateral opacities on imaging consistent with pulmonary edema and hypoxia as defined of PaO2/FiO2 ratio of <300.

Other variables of interest

The following data on severe sepsis/septic shock patients was collected: demographics, comorbid conditions, medication history, glucose control during ICU stay, insulin administration, ventilator settings (e.g. tidal volume,

compliance), severity of illness measures (APACHE, SOFA, lung injury score), duration of mechanical ventilation, development of ARDS, survival status. Data was collected until death or patient discharge.

General methods

Sample collection and processing

A one-time blood draw and bronchoscopy with bronchoalveolar lavage (BAL) were performed within 72 hours of sepsis development and within 24 hours of enrollment into the study. Experienced staff participating in the study performed the bronchoscopy.[29] For the BAL a total of 180 ml of normal saline were instilled. For the blood draw a standard 4ml lavender blood collection tube was used.

BAL and plasma processing: Plasma and BALF were processed same day, within 2h of collection. The bronchoalveolar lavage fluid was centrifuged and once the cells were separated from the BAL, the fluid was retained and stored for metabolomic analysis at -80°C. Plasma was processed in a similar fashion. For the metabolomics analysis 100 µl aliquots of plasma and BAL samples were treated with acetonitrile, spiked with internal standard mix, and centrifuged to remove protein prior to loading on Shimadzu auto sampler.

Data management and statistics

Power calculations and sample size:

The study is explorative and hypothesis generating. In a previous study using plasma from ARDS patients, significant metabolic differences were seen with a sample size of 13 ARDS patients [28], another study was able to identify differences in a total of 28 patients[26]. Based on these previous studies, we enrolled 79 severe sepsis/septic shock patients.

Data analysis and statistical methods:

Data was collected using established case report forms (CRFs) and then later entered into a web-based database. Univariate comparisons between patients who did and did not develop ARDS and between survivors and non-survivors was calculated and evaluated for statistical significance at an alpha of 0.05 using a chi-squared test for categorical variables and a two-sample t-test for continuous variables. Wilcoxon Rank Sum test was used for variables not normally distributed.

Metabolomic analysis was performed using the statistical software R (version 3.0.0). Samples were analyzed using the Q Exactive mass spectrometer in a positive mode. After raw data extraction multiple steps of data normalization and data reduction were performed prior to further analysis. We used the technical average of each sample (each sample was measured 3 times and we took the mean). Only features of at least 70% signal in either of the two groups (survivors vs. non-survivors and ARDS vs. non-ARDS) were retained for further analysis. Then we performed log2 transformation to reduce heteroscedasticity,

median centering of metabolites and quantile normalization were also performed.

At each log2 change filter we performed paired t-tests to identify differentiating metabolites with a false discovery rate of ≤10% to account for multiple comparisons.[30] Then two-way hierarchical clustering analysis was used for visualization and identification of global patterns.[31] This was followed by k-fold cross-validation and determination of the optimization score to determine results for downstream analysis.

After we identified the candidate differentiating metabolites for both outcomes (ARDS yes vs. no and survivors vs. non-survivors) we used metabolome-wide Spearman correlation network analysis to identify which pathways were involved. Using existing approaches, such as the Metlin database, we mapped the top 500 metabolites identified in the network analysis to specific pathways.[32]

The software SAS (version 9.2) was used to perform multivariate logistic regression to identify which metabolite clusters were significantly associated with the two outcomes of interest (28-day mortality and development of ARDS). We took the clusters identified in the two-way hierarchical analysis (where the cluster variable in the regression represents the average metabolite intensity within that cluster) and regressed them over the outcome of interest (death and ARDS). The other variables included in the two models were gender, age and severity of illness score (APACHE score). Stepwise elimination method was used to select the final model using $p<0.1$ for variables to entry and stay in the model.

RESULTS

Enrollment began in December 2009 and was completed in September 2012. BAL and plasma samples were obtained from a total of 79 patients. Forty-two of 79 developed ARDS, the mean age of the cohort was 53.9 years, 59 % were male, and 68% were black (Table 3a). The overall mortality for the cohort was 48%.

When looking at the development of ARDS (Table 3a), there was no mortality difference between ARDS and non-ARDS patients (42.9% vs. 54.1%, p=0.32). There was no difference in APACHE score, ICU and hospital length of stay, ventilator days. Interestingly, ARDS patients developed significantly less shock (50% vs. 72%, p=0.048). Also, ARDS patients had significantly more bloodstream infections (86% vs. 49%, p=0.0004) and significantly less GU infections(0% vs. 24%) compared to non-ARDS patients.

When looking at the other outcome, 28-day mortality, (Table 3b), there was no difference between survivors and non-survivors in terms of PF ratio, development of ARDS, type of infection implicated in the development of sepsis, ICU and hospital length of stay, but as expected, the survivors had a significantly lower APACHE (severity of illness) score.

Metabolites BAL fluid:

A total of 9,148 metabolites were identified in the BAL fluid. After statistical selection, when looking at survivors vs. non survivors we found 37 differentially

expressed metabolites (DEMs) in 11 clusters, with 85.9% 10-fold cross-validation classification accuracy and FDR of 10%. The two-way hierchical clustering is displayed in **Figure 1.** helps us visualize the differential metabolite expression where the rows are comprised of metabolites (clustered based on level of similarity in their pattern of intensity) and the survivors vs. non-survivors are displayed in the columns (clustered based on their similarities in overall metabolic pattern). The color key signifies the level of metabolite intensity (ranging from blue- negative expression to yellow- positive expression)

Figure 2. shows the differential expression of individual metabolites within the clusters. One can see that in clusters 1, 3, 5 and 9 the average metabolite intensity was lower in non-survivors vs. survivors where as in clusters 2, 4, 6, 7, 8, 10 and 11 the average intensity of metabolites in non-survivors was higher compared to survivors.

Multivariate logistic regression was performed with the identified 11 clusters and showed only clusters 1, 2 and 10 were significantly associated with the outcome death, and clusters 2 and 10 were associated with a higher risk of death (OR 1.9 and 1.2 respectively), whereas cluster 1 was associated with a lower risk of death (Table 4). Every unit increase in APACHE score was also significantly associated with higher risk of death (OR 1.9). Age and gender were not significantly associated.

The 37 DEMs were then included in metabolome network analysis looking for metabolite associations that would give a better understanding of the involved pathways. The top 500 metabolites from the network analysis (using METLIN)

were found to be significantly enriched in the ubiquinone and other terpenoidquinone biosynthesis (vitamin and cofactor biosynthesis); glycerophospholipid metabolism; drug metabolism (cytochrome p450); and valine, leucine and isoleucine biosynthesis.

BALF analysis between **ARDS and non-ARDS patients** showed 35 DEMs in 20 clusters with 92.1% cross-validation accuracy and FDR of 10%.

Figure 3. helps us visualize the differential metabolite expression where the rows are comprised of metabolites (clustered based on level of similarity in their pattern of intensity) and the ARDS vs. non-ARDS samples are displayed in the columns (clustered based on their similarities in overall metabolic pattern). The color key signifies the level of metabolite intensity (ranging from blue- negative expression to yellow- positive expression). **Figure 4.** shows the differential expression of individual metabolites within the 20 clusters where ARDS patients had an average lower intensity of metabolite expression in clusters 1,8,10,11,13,16,18,and 19 and overall higher average intensity in the rest of the clusters compared to non-ARDS patients.

Multivariate logistic regression was performed with the identified 20 clusters and showed only clusters 1, 2, 6,11 and 13 were significantly associated with the outcome death, and clusters 2 and 6 were associated with a higher risk of developing ARDS (OR 1.2 and 1.5 respectively), whereas clusters 1, 11 and 13 ere associated with a lower risk of developing ARDS (Table 5). Age, gender and

APACHE score were not significantly associated with the development of ARDS.

When the 35 differentially expressed metabolites were used in the network analysis the following pathways were implicated: phenylalanine metabolism, tyrosine, D-arginine, D-ornithine, drug metabolism (cytochrome p450).

Plasma Metabolites:

Analysis of plasma discovered 7,313 metabolites but none were differentially expressed in regards to development of ARDS or mortality after FDR correction.

DISCUSSION

The results of the study show that sepsis-induced ARDS causes measurable metabolite changes that can be captured for use as a diagnostic and prognostication tool and for the identification of potential novel biomarkers of ARDS. Previous studies have compared ARDS patients and healthy controls, [25] but to our knowledge this is the first quantitative metabolomics study that compares sepsis patients in regards to development of ARDS and 28-day mortality. We were able to identify metabolic patterns that distinguish sepsis patients early in their hospital course in regards to mortality and development of ARDS. Interestingly, one pathway that was involved in both outcomes was drug metabolism (cytochrome p450), which may be a clue that differences in medication metabolism, and maybe specifically antibiotics, may be linked to mortality and/or development of ARDS.

Interestingly, a lot of the previous studies have used plasma as the main source of metabolite information [25, 33] because plasma is much more readily available and easier to obtain. Because it interacts with all tissues, one could suggest that it represents the physiological 'average' of the organism's biochemical information. Changes related to sepsis, a systemic process, should be readily identifiable in plasma. However, we were only able to identify differences in metabolic composition in our two outcomes of interest in the BALF only, which suggests that the pathophysiologic process that occurs in the lungs leading to ARDS and mortality is not necessarily reflected in the plasma. A likely

explanation is that although the alveolar-capillary barrier is damaged and leaky, it leads to influx of cells and fluid into the alveoli but not back into the bloodstream and sampling of the lungs might be necessary to determine diagnosis and prognosis. One can argue that sampling of the BAL fluid in critically ill patients is too invasive and not always feasible. There has been promising research that mini-BAL can be used instead [34, 35], which does not require a bronchoscopy and is instead obtained at the bedside using the small plastic catheter that can be inserted through the endotracheal tube without a bronchoscope.

The strengths of our study are that it includes a relatively high number of patients (n=79) compared to previous studies, and all subjects were critically ill and at risk of developing ARDS. This makes the results we produced more useful in real-life settings. Another strength is the use of metabolomics is that its objective approach of discovery that can potentially reveal new pathways that have not been implicated before. While the analysis is not yet mainstream it can be performed in a relatively inexpensive and rapid way. Metabolomics also describes the downstream function of protein alteration and thus presents the true phenotype of a disease process.

The limitations of the study are that all patients were recruited from a single center, and they were predominantly male and black. Thus, the results may not be generalizable to other centers and patient demographics. Also, while all samples were obtained within 72 hours of admission to the ICU and the development of sepsis, there is some variation of the timing, which may have

impacted the results by catching patients at various times within the disease process. A possible future project would involve serial sampling; possibly daily for the first seven days to see how the metabolite profile changes over time in an individual patient and whether this correlates to outcomes. It would likely help us more precisely narrow down the pathways involved. A limitation of metabolomics is that it is relatively new and requires skilled personnel for both measurement and interpretation since the analysis can be very complex. No normal ranges have been established and processes other than the disease, such as nutrition and medications, may affect the metabolites. There is also limited data on whether the results are reproducible. All biological samples collected for metabolic analysis require careful sample handling and they usually stored at -80°C but here is no data on metabolite stability during storage.

The results of this study show that there is a potential to use metabolomics as a quick diagnostic tool to identify patients who are more likely to die or develop ARDS as soon as a few hours of admission, which will help guide clinical care and prognostication. A more pivotal goal is to identify the pathways involved and use basic science approach to see what type of intervention on molecular or protein level would prove beneficial to stop or reverse the processes leading to adverse outcomes.

Essentially, the field of metabolomics is well positioned to play a critical role in biomarker discovery in patients with critical illness but a lot more work has yet to be done. First the normal human metabolome needs to be defined; more studies are needed to quantitatively characterize the normal blood

metabolome to precisely describe the "normal range" of the multiple endogenous metabolites in blood. There is also no standard for analysis and interpretation and no single technique currently exists that can identify all the metabolites in the metabolome.

Similarly to the Human Genome Project, The Human Metabolome Project (http://www.hmdb.ca) is compiling reference range data and developing bioinformatics tools to assist in mapping the human metabolome which will help greatly. So, although the field of metabolomics holds great promise for advancing to translational research, there are still a lot of gaps in our knowledge that need to be addressed before it's full potential can be realized in ARDS and beyond.

Table 1. Comparison of the American-European Consensus Conference definition and Berlin definition of Acute Lung Injury (ALI)/ Acute Respiratory distress syndrome (ARDS)

Characteristic	The AECC definition 1994	The Berlin definition 2012
Onset	Acute	Within 7 days of predisposing clinical insult
Radiographic abnormality	Bilateral infiltrates on frontal chest radiograph	Bilateral opacities on CXR or CT, not fully explained by effusion, atelectasis, or nodules
Noncardiogenic source of pulmonary edema	No clinical evidence of elevated left atrial pressure or PCWP of ≤ 18 mmHg	Respiratory failure not fully explained by cardiogenic pulmonary edema or volume overload
Oxygenation	PaO2/FiO2 ratio ALI<300 ARDS<200	PaO2/FiO2 ratio with 5 cmH2O positive end-expiratory pressure mild ARDS 201-300, moderate ARDS 101-200, severe ARDS <100
Predisposing condition	Not specified	If none identified need to rule out cardiogenic edema

Table 2. Predisposing conditions associated with Acute Respiratory Distress syndrome (ARDS)

Direct lung injury	Indirect lung injury
Pneumonia	Sepsis
Aspiration	Blood transfusion
Lung contusion	Trauma
Toxic inhalation	Pancreatitis
Near-drowning	Burn

Table 3a. Demographics, clinical and physiological characteristics ARDS vs. non-ARDS patients

Variable	ARDS (N=42)	Non-ARDS (N=37)	
		N	P-value
Age (mean ±SD)	55 (±15)	55 (±16)	0.98
Sex			0.98
- male 26	21		
- female 16	16		
Race			0.98
- **white**	5	4	
- **black**	36	32	
- **other**	1	1	
Mortality %	43	54	0.32
Shock (%)	50	72	0.048*
APACHE score (mean, ±SD)	23 (±7)	26(±6.9)	0.057
BMI (median, IQR)	25(22-26)	22(18-27)	0.09
Tobacco use %	63	65	0.95
Vent days (median, IQR)	11(6-18)	11(6-17)	0.97
ICU LOS (median, IQR)	14 (6-23)	14 (10-26)	0.59
Hospital LOS (median, IQR)	25 (5-33)	20(13-22)	0.47
Source of infection (%)			
- **Respiratory**	12	14	NS
- **Blood**	86	49	0.0004*
- **GU**	0	24	0.0007*
- **CSF**	0	3	NS
- **GI**	0	5	NS
- **Other**	0	5	NS

α=0.05, significant p values are marked with an asterisk(*)
NS= non-significant
LOS= length of stay (ICU, hospital)
APACHE= Acute Physiologic and Chronic Health Evaluation

Table 3b. Demographics, clinical and physiologic characteristics survivors vs. non-survivors

Variable	Survivors(N=41) N	Non-Suvivors (N=38) N	P-value
Age (±SD)	52.2(±15.6)	57.2(±16.2)	0.17
Sex			0.78
- male	25	22	
- female	16	16	
Race			0.07
- **white**	2	7	
- **black**	37	31	
- **other**	2	0	
ARDS	24	18	0.32
PF ratio (±SD)	148(±68)	161(±84)	0.58
APACHE score (±SD)	21.9(±5.88	27.8(±6.9)	0.0001*
BMI	25.2	23.9	0.5
Tobacco use	25	25	0.41
Vent days (±SD)	15.1(±27.7)	22.6(±33.8)	0.288
ICU LOS (±SD)	20.1(±28.3)	23.7(±32.4)	0.59
Hospital LOS (±SD)	30.3 (±27.8)	29.4(±33.6)	0.89
Source of infection			0.47
- **Respiratory**	31	23	
- **Blood**	5	5	
- **GU**	4	5	
- **CSF**	0	2	
- **GI**	0	1	
- **Other**	1	2	

α=0.05, significant p values are marked with an asterisk(*)
NS= non-significant
LOS= length of stay (ICU, hospital)
APACHE= Acute Physiologic and Chronic Health Evaluation

Figure 1. Two-way hierarchical clustering analysis in survivors vs. non-survivors. On the horizontal axis are the samples, color-coded. On the vertical axis are the metabolites. Both patients and metabolites are clustered based on similarity. The metabolite expression is coded from blue (negative) to yellow (positive).

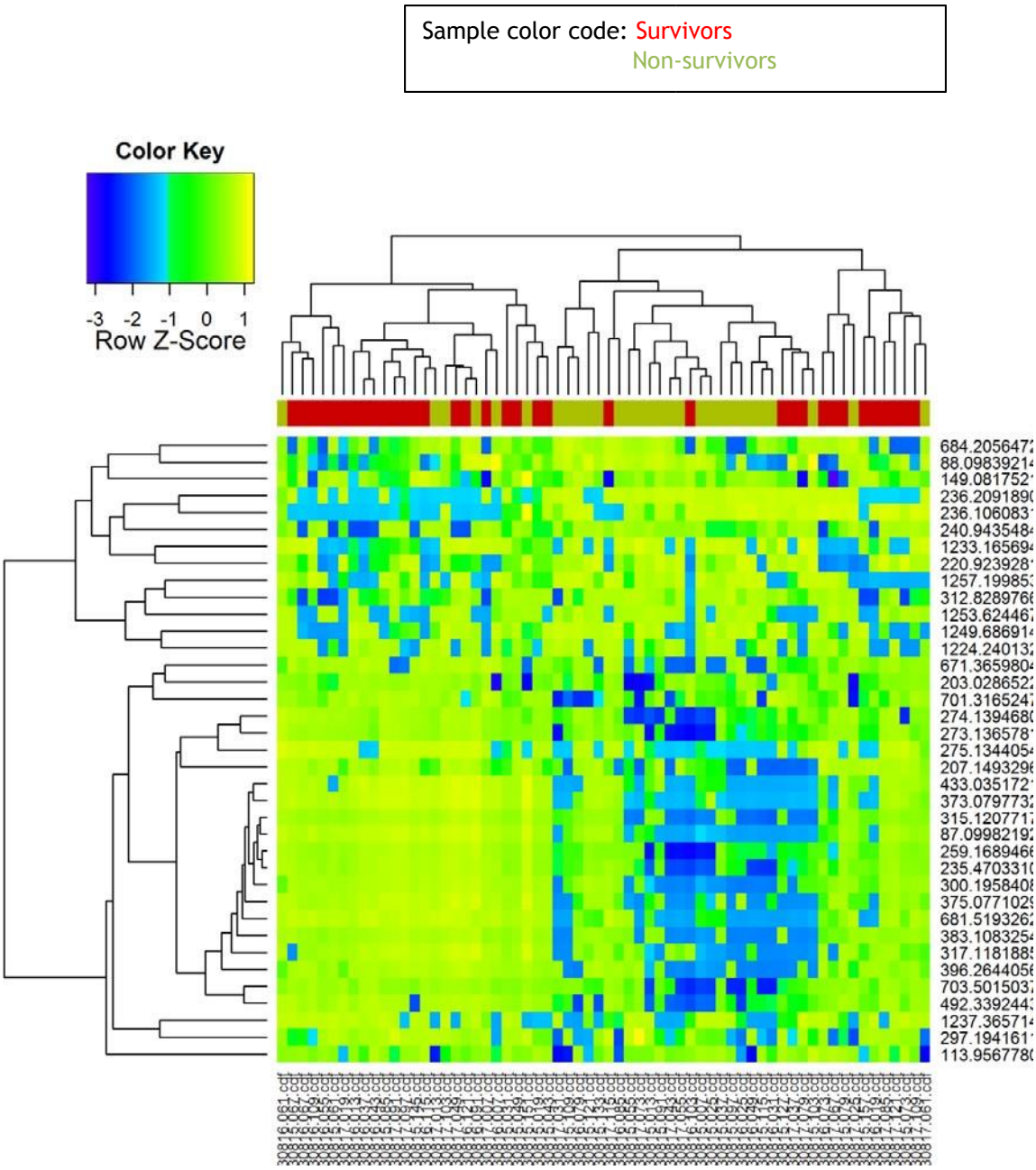
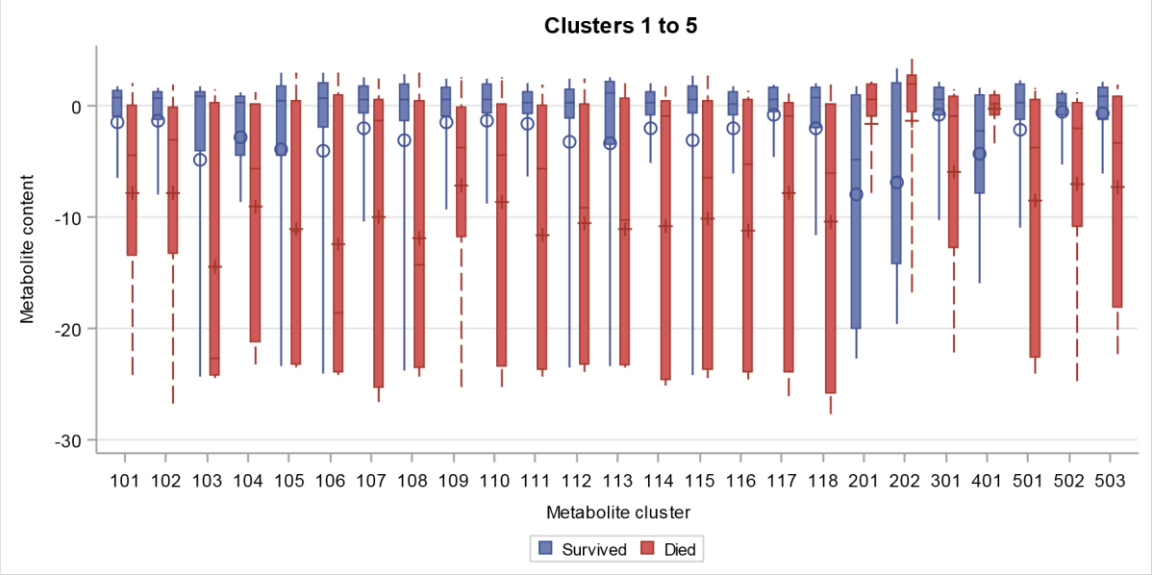


Figure 2. Boxplot analysis showing differential expression of metabolites in survivors vs. non-survivors within individual clusters. Metabolites 101-118 represent the 18 metabolites within cluster 1, metabolites 201 nd 202 are the two metabolites within cluster 2, etc.



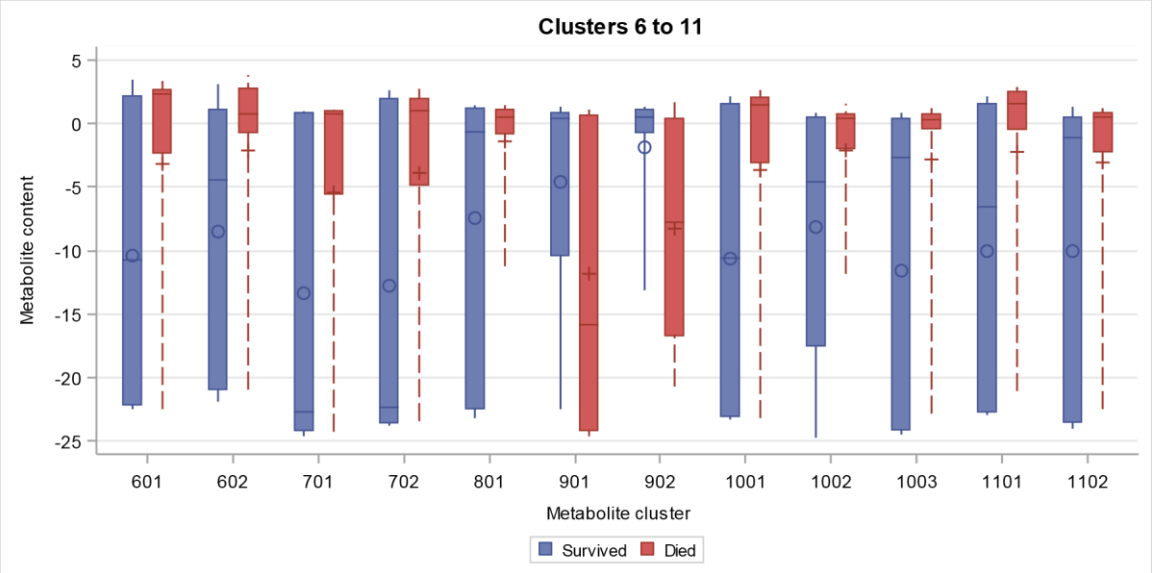


Table 4. Final model of multivariate logistic regression using the 37 differentially expressed metabolites in 11 clusters, with death as the outcome. Only clusters 1,2 and 10, together with the severity of illness score (APACHE) were found to be significantly associated.

	β	SE	OR	95% CI	P
Cluster 1	-0.2715	0.1094	0.762	0.615,0.944	0.013
Cluster 2	0.6462	0.2567	1.908	1.154,3.156	0.01
Cluster 10	0.1911	0.0769	1.211	1.041,1.408	0.013
Age (per 1 yr increase)	0.0337	0.0412	1.034	0.954,1.121	0.41
APACHE (per 1 unit increase)	0.6264	0.2626	1.871	1.118,3.13	0.017

Figure 3. Two-way hierarchical clustering analysis ARDS vs. non-ARDS. On the horizontal axis are the samples, color-coded. On the vertical axis are the metabolites. Both patients and metabolites are clustered based on similarity. The metabolite expression is coded from blue (negative) to yellow (positive).

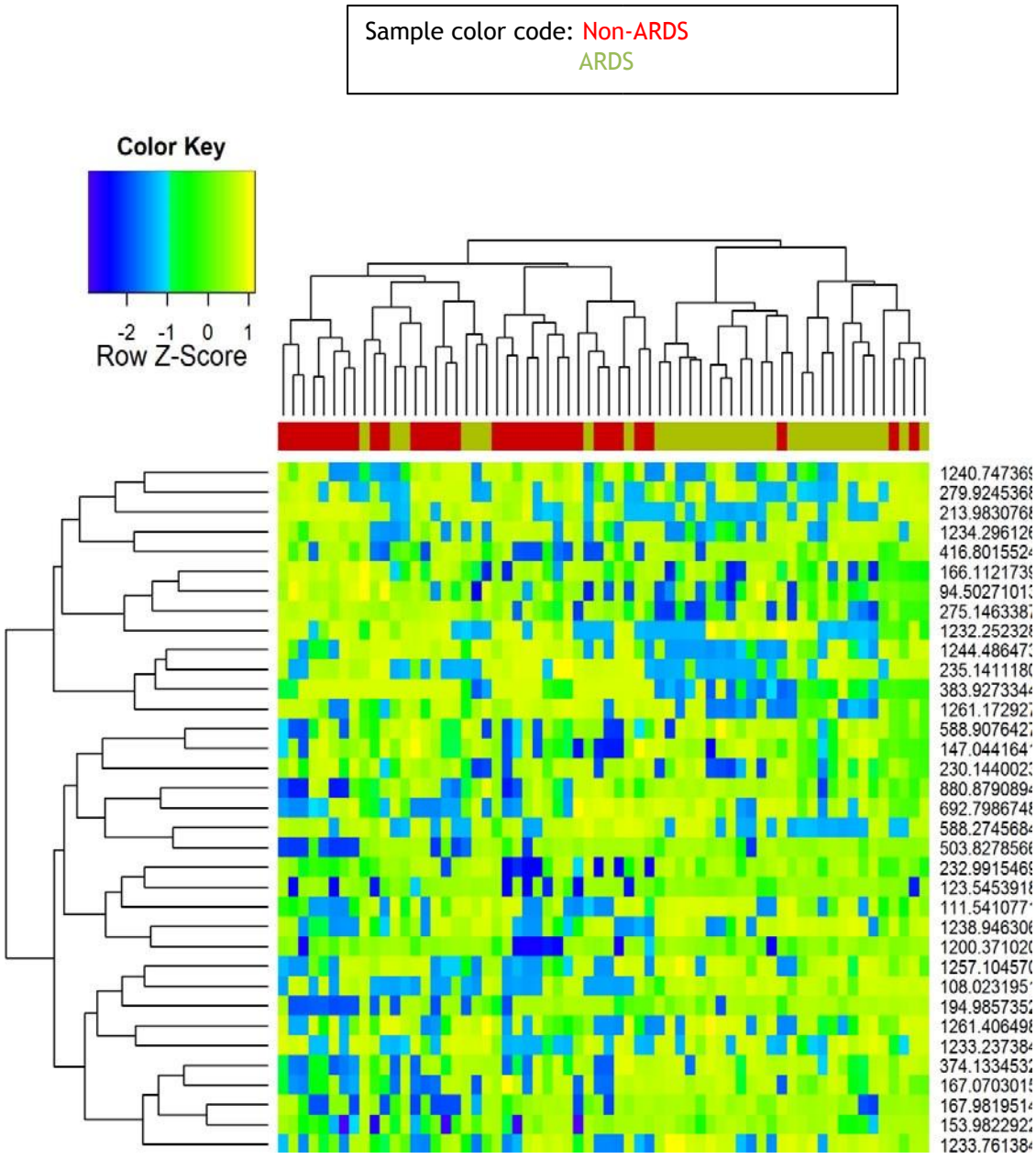


Figure 4. Boxplot analysis showing differential expression of metabolites ARDS vs. non-ARDS

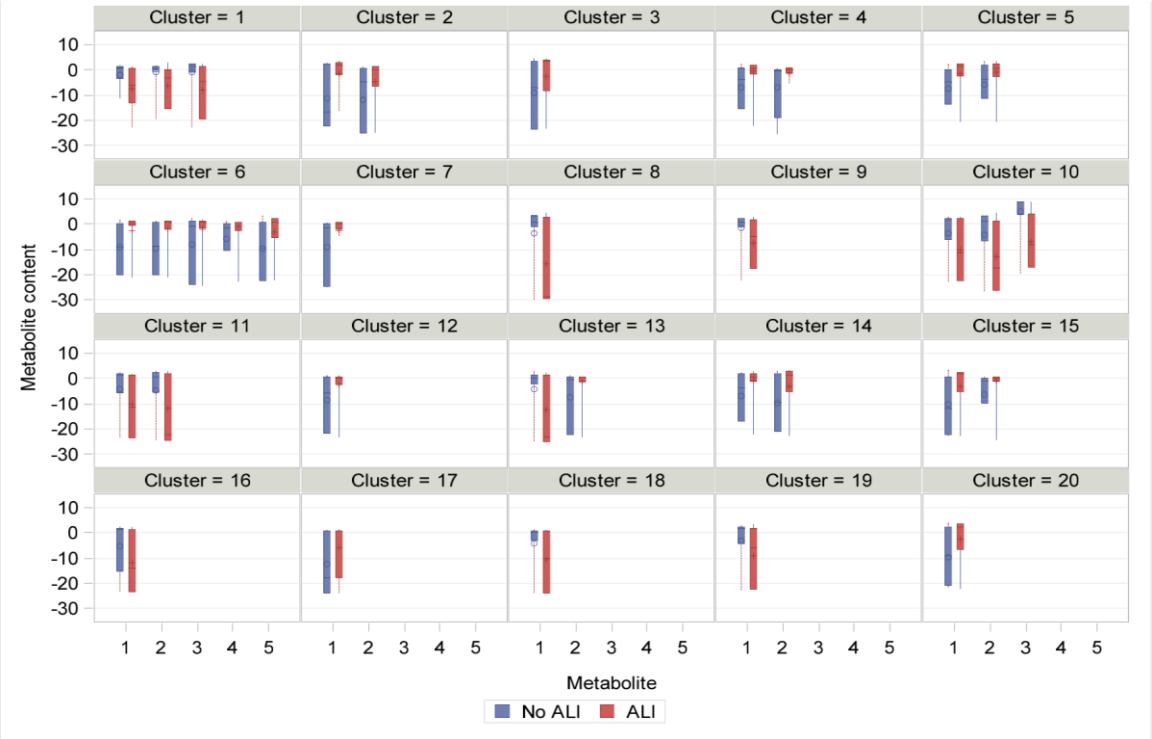


Table 5. Final model of multivariate logistic regression using the identified 35 metabolites in 20 clusters, with ARDS as the outcome. Only clusters 1,2,6,11, and13 were found to be significantly associated with ARDS

	β	SE	OR	95% CI	P
Cluster 1	-0.5705	0.2091	0.565	0.375 , 0.852	0.006
Cluster 2	0.2207	0.1023	1.247	1.02 , 1.524	0.03
Cluster 6	0.4246	0.1864	1.529	1.061 , 2.203	0.02
Cluster 11	-0.3282	0.1157	0.72	0.574 , 0.904	0.005
Cluster 13	-0.3052	0.1318	0.737	0.569 , 0.954	0.02
Gender (0 vs 1)	0.0409	1.3157	1.042	0.079 , 13.732	0.98
Age (per 1 yr increase)	-0.00968	0.0449	0.99	0.907 , 1.082	0.83
APACHE (per 1 unit increase)	-0.0518	0.1033	0.949	0.776 , 1.162	0.62

References:

1. Ashbaugh, D.G., et al., *Acute respiratory distress in adults.* Lancet, 1967. **2**(7511): p. 319-23.
2. Walkey AJ, S.R., Ho V, Alkan P, *Acute respiratory distress syndrome: epidemiology and management approaches.* Clin Epidemiol, 2012(4): p. 159–169.
3. Ferguson, N.D., et al., *The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material.* Intensive Care Med, 2012. **38**(10): p. 1573-82.
4. Piantadosi, C.A. and D.A. Schwartz, *The acute respiratory distress syndrome.* Ann Intern Med, 2004. **141**(6): p. 460-70.
5. Jr, T.J., *Pulmonary pathology of the adult respiratory distress syndrome.* Clin Chest Med, 1990. **11(4)**.
6. Hudson, L.D., et al., *Clinical risks for development of the acute respiratory distress syndrome.* Am J Respir Crit Care Med, 1995. **151**(2 Pt 1): p. 293-301.
7. Moss, M., et al., *The role of chronic alcohol abuse in the development of acute respiratory distress syndrome in adults.* JAMA, 1996. **275**(1): p. 50-4.
8. MacCallum, N.S. and T.W. Evans, *Epidemiology of acute lung injury.* Curr Opin Crit Care, 2005. **11**(1): p. 43-9.
9. Rubenfeld, G.D., et al., *Incidence and outcomes of acute lung injury.* N Engl J Med, 2005. **353**(16): p. 1685-93.
10. Villar, J., et al., *The ALIEN study: incidence and outcome of acute respiratory distress syndrome in the era of lung protective ventilation.* Intensive Care Med, 2011. **37**(12): p. 1932-41.
11. Calfee, C.S., et al., *Use of risk reclassification with multiple biomarkers improves mortality prediction in acute lung injury.* Crit Care Med, 2011. **39**(4): p. 711-7.
12. Binnie, A., J.L. Tsang, and C.C. dos Santos, *Biomarkers in acute respiratory distress syndrome.* Curr Opin Crit Care, 2014. **20**(1): p. 47-55.
13. Annane, D., E. Bellissant, and J.M. Cavaillon, *Septic shock.* Lancet, 2005. **365**(9453): p. 63-78.
14. Walkey AJ, W.R., Lindenauer PK, *Utilization patterns and outcomes associated with central venous catheter in septic shock: a populationbased study.* Crit Care Med, 2013. **41(6)**.
15. Danai, P. and G.S. Martin, *Epidemiology of sepsis: recent advances.* Curr Infect Dis Rep, 2005. **7**(5): p. 329-34.
16. Dellinger, R.P., et al., *Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012.* Intensive Care Med, 2013. **39**(2): p. 165-228.

17. Bastarache, J.A., L.B. Ware, and G.R. Bernard, *The role of the coagulation cascade in the continuum of sepsis and acute lung injury and acute respiratory distress syndrome*. Semin Respir Crit Care Med, 2006. **27**(4): p. 365-76.
18. Ware LB, M.M., *The acute respiratory distress syndrome*. N Engl J Med 2000. **342 (18)**: p. 1334-1349

.

19. Biomarkers Definitions Working, G., *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework*. Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.
20. Safety, W.I.P.o.C., *Biomarkers in risk assessment: validity and validation*. 2001.
21. Janz, D.R. and L.B. Ware, *Biomarkers of ALI/ARDS: pathogenesis, discovery, and relevance to clinical trials*. Semin Respir Crit Care Med, 2013. **34**(4): p. 537-48.
22. Griffin, J.L., *Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis*. Curr Opin Chem Biol, 2003. **7**(5): p. 648-54.
23. Beckonert, O., et al., *Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts*. Nat Protoc, 2007. **2**(11): p. 2692-703.
24. Angela J. Frank, e.a., *Metabolomic Analysis Of Exhaled Breath Condensate In Acute Respiratory Distress Syndrome*. Am. J. Respir. Crit. Care Med. , 2011. **183**.
25. Stringer, K.A., et al., *Metabolic consequences of sepsis-induced acute lung injury revealed by plasma (1)H-nuclear magnetic resonance quantitative metabolomics and computational analysis*. Am J Physiol Lung Cell Mol Physiol, 2011. **300**(1): p. L4-L11.
26. Jose L. Izquierdo-Garcia, P., Jesús Ruiz-Cabello, PhD , Pablo Cardinal, MD , Pilar Fernandez-Segoviano, MD, Andrés Esteban, MD , José A. Lorente, MD PhD, *Metabolomic Analysis As A Diagnostic Tool For Acute Respiratory Distress Syndrome Caused By H1N1 Influenza Infection In Humans*. Am Journal of Resp and Crit Care Med, 2012. **185**.
27. A. M. Esper, Y.P., G. S. Martin, D. Jones, *Metabolomic Analysis In Sepsis And ARDS*. Am J Respir Crit Care Med 2013.
28. Park, Y., et al., *Metabolic effects of albumin therapy in acute lung injury measured by proton nuclear magnetic resonance spectroscopy of plasma: a pilot study*. Crit Care Med, 2011. **39**(10): p. 2308-13.
29. Steinberg, K.P., et al., *Safety of bronchoalveolar lavage in patients with adult respiratory distress syndrome*. Am Rev Respir Dis, 1993. **148**(3): p. 556-61.
30. Storey, J.D., Tibshirani, R., *Statistical significance for genomewide studies*. Proc Natl Acad Sci U S A, 2003. **100(16)**: p. 289-300.

31. Benjamini, Y.H.Y., *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* J Roy Statist Soc Ser B (Methodological) 1995. **57**: p. 289-300.
32. Kanehisa, M., *The KEGG database.* Novartis Found Symp, 2002(247): p. 91-101.
33. Serkova, N.J., T.J. Standiford, and K.A. Stringer, *The emerging field of quantitative blood metabolomics for biomarker discovery in critical illnesses.* Am J Respir Crit Care Med, 2011. **184**(6): p. 647-55.
34. Rai, R., *Metabolomic profiling in human lung injuries by high resolution nuclear magnetic resonance spectroscopy of BAL fluid.* Metabolomics, 2013(3): p. 667 – 676.
35. Singh, C., et al., *Mini-bronchoalveolar lavage fluid can be used for biomarker identification in patients with lung injury by employing (1)H NMR spectroscopy.* Crit Care, 2013. **17**(2): p. 430.