Host-Response Subphenotypic Classification with A Parsimonious Model Offers Prognostic Information in Patients with Acute Respiratory Failure: A Prospective Cohort Study.

Callie M. Drohan
UPMC Presbyterian https://orcid.org/0000-0003-2578-0413
S. Mehdi Nouraie
University of Pittsburgh Medical Center
William Bain
University of Pittsburgh Medical Center
Faraaz A. Shah
University of Pittsburgh Medical Center
John Evankovich
Columbia University Medical Center
Yingze Zhang
University of Pittsburgh Medical Center
Alison Morris
University of Pittsburgh Medical Center
Bryan J. McVerry
University of Pittsburgh Medical Center
Georgios D. Kitsios (kitsiosg@upmc.edu)
University of Pittsburgh Medical Center

Research

Keywords: Subphenotypes, Acute Respiratory Distress Syndrome, Latent Class Analysis, heterogeneity, mechanical ventilation

Posted Date: August 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-57907/v1

Abstract

Background: Recent research in patients with ARDS has consistently shown the presence of two distinct subphenotypes of host-responses (hyper- and hypo-inflammatory) with markedly different outcomes and responses to therapies. However, inherent uncertainty in reaching the diagnosis of ARDS creates considerable biological and clinical overlap with other broadly-defined syndromes of acute respiratory failure, such as patients with risk factors (e.g. sepsis or pneumonia) for ARDS (at-risk for ARDS [ARFA]) or patients with decompensated congestive heart failure (CHF). Limited data are available for the presence of subphenotypes in such broader critically-ill populations.

Methods: We enrolled mechanically-ventilated patients with acute respiratory failure (ARDS, ARFA, and CHF) and measured 11 plasma biomarkers at baseline. We applied latent class analysis (LCA) methods to determine optimal subphenotypic classifications in this inclusive patient cohort by considering clinical variables and biomarkers. We then derived a parsimonious logistic regression model for subphenotype predictions and compared clinical outcomes between subphenotypes.

Results: We included 334 patients (123 [37%] ARDS, 177 [53%] ARFA, 34 [10%] CHF) in a derivation cohort and 36 patients in a temporally-independent validation cohort. A two-class LCA model was found to be optimal, classifying 29% of patients in the hyper-inflammatory subphenotype, consistent with prior findings. A 4-variable parsimonious model (angiopoietin-2, soluble tumor necrosis factor receptor-1, procalcitonin and bicarbonate) for subphenotype prediction offered excellent classification (area under the curve = 0.98) compared to LCA classifications. For both LCA- and regression model classifications, hyper-inflammatory patients had higher severity of illness by Sequential Organ Failure Assessment scores, fewer ventilator-free days and higher 30- and 90-day mortality (all $p<0.01$) compared to the hypo-inflammatory group. Subphenotype predictions in the validation cohort revealed consistent trends for clinical outcomes and higher levels of inflammatory biomarkers in the hyper-inflammatory group (22%).

Conclusions: Host-response subphenotypes are observable in broader and heterogeneous patient populations beyond just patients with ARDS, and subphenotypic classifications offer prognostic information on clinical outcomes. Accurate subphenotyping is possible with the use of a simple predictive model to improve clinical applicability.

Introduction:

The substantial biological and clinical heterogeneity of clinical syndromes in critical care, such as ARDS, sepsis or pneumonia, have made it difficult to identify effective therapeutic interventions for individual patients (1–4). There is growing interest to identify specific subgroups of critically-ill patients with distinct biological mechanisms, so that targeted treatments can be delivered. Prior studies in patients with ARDS have consistently shown the presence of two distinct subgroups (hyper- and hypo-inflammatory subphenotypes) with markedly different clinical outcomes and responses to therapies, such as positive end-expiratory pressure levels, statin use, and fluid management strategies (5). These

subphenotypes were initially discovered from independent, unsupervised examinations of randomized clinical trial data with latent class analysis (LCA) of both clinical and biomarker variables. However, LCA and other unsupervised clustering approaches are laborious and cannot be implemented clinically. For that reason, efforts have been made to derive simpler models for subphenotype prediction in patients with ARDS, involving either predictive models with biomarkers (6) or machine-learning classifiers based on clinical variables alone (7).

We recently discovered that the two distinct subphenotypes of host-responses are also present in patients with risk factors for ARDS (e.g. pneumonia or sepsis) who do not meet the clinical criteria of ARDS diagnosis (i.e. bilateral opacities on chest-radiographs and severe hypoxemia) (8). In such patients at-risk for ARDS (ARFA), classification to a hyper-inflammatory subphenotype by LCA was consistently associated with higher severity of illness, persistently elevated biomarkers of host injury and inflammation, and worse clinical outcomes. These observations suggest that common pathways of systemic inflammatory injury are likely present in broader critically-ill populations, beyond the subset of patients with ARDS (~ 10% of all ICU admissions) (2). Additionally, inherent subjectivity in making the diagnosis of ARDS, primarily due to clinicians' disagreement on radiographic criteria or exclusion of cardiac failure in the Berlin definition, is likely to generate considerable overlap among patient subgroups diagnosed as ARDS, ARFA or with cardiogenic pulmonary edema from decompensated congestive heart failure (CHF) (9–12). Therefore, subphenotyping efforts in a more inclusive diagnostic framework that considers broader patient populations with acute respiratory failure (ARDS, ARFA or CHF) may uncover distinct subgroups of patients that may benefit from targeted enrollment in clinical trials. We sought to determine whether hyper- and hypo-inflammatory subphenotypes could be detected in a heterogeneous critically-ill population with acute hypoxemic respiratory failure, and whether a parsimonious predictive model with a small number of relevant variables could be derived for future clinical or investigative application.

Methods:

Extensive methods are provided in Additional file 1.

Clinical cohort:

From October 2011 to August 2019 we prospectively enrolled a convenience sample of consecutively admitted mechanically-ventilated patients with acute respiratory failure in Medical Intensive Care Units (ICUs) at the University of Pittsburgh Medical Center to the Pittsburgh Acute Lung Injury Registry (ALIR) and Biospecimen Repository(8, 13–15). We excluded patients unable to provide informed consent or mechanical ventilation for > 72 hours prior to enrollment. Written informed consent was provided by all participants or their surrogates. The study was approved by the University of Pittsburgh Institutional Review Board (protocol STUDY19050099). We recorded baseline demographics, comorbidities, mechanical ventilation and laboratory variables, and calculated sequential organ failure assessment (SOFA) scores.

Biomarker measurements:

From enrolled subjects, we collected blood samples within 48hr of intubation. We measured 10 host-response biomarkers that have been shown to have validated associations with ARDS and/or sepsis with a customized Luminex assay (R&D Systems, Minneapolis) (16), as previously described. These host-response biomarkers were classified into those measuring innate immune responses (interleukin [IL]-6, IL-8, IL-10, fractalkine, soluble tumor necrosis factor receptor-1 [TNFR1], suppression of tumorigenicity-2 [ST-2]) (6, 17–20), epithelial injury (receptor of advanced glycation end-products [RAGE]) (21, 22), endothelial injury (angiopoietin-2 [ang-2]) (23–25), and response to bacterial infections (procalcitonin and pentraxin-3) (26–28). We also measured 1-3-beta-D-glucan (BDG), a fungal cell-wall constituent shown to correlate with inflammatory biomarkers, using the commercially available Fungitell® Limulus Amebocyte Lysate (LAL) assay (Associates of Cape Cod, Inc, East Falmouth, MA, USA) at the manufacturer's facility (29).

Clinical group classifications:

A consensus committee retrospectively reviewed all available clinical and radiographic data without knowledge of biomarkers values and classified subjects into distinct clinical categories of acute respiratory failure: a) ARDS per Berlin criteria (10), b) (ARFA) based on presence of an identifiable lung injury risk factor upon enrollment but not fulfilling ARDS criteria, c) cardiogenic pulmonary edema from CHF, d) acute on chronic respiratory failure (e.g. acute exacerbation of chronic obstructive pulmonary disease [COPD]), e) intubation for airway protection, and f) "other" category, including cases for which the committee could not reach consensus for clinical classification into any of the categories above.

Outcomes:

Primary outcomes included ventilator-free days (VFD) and 30- and 90-day mortality. Patients were also followed prospectively for incidence of shock within the first week of enrollment (defined as need for vasopressor agents), acute kidney injury (AKI), time-to-liberation from mechanical ventilation, and ICU length of stay.

Subphenotypic classifications and statistical analyses

For subphenotypic classification with LCA models first and then derivation of a parsimonious predictive model, we considered patients from the ARDS, ARFA, and CHF groups in our primary analyses. Our selection of these three primary clinical groups was made on the basis of biological similarity (similar distribution of subphenotypes between ARDS and ARFA patients in our previous study) (8), as well as due to the higher clinical relevance of risk stratification and treatment selection in these critically-ill patients, compared to patients intubated for airway protection or acute exacerbation of COPD for example.

For our primary analysis, we divided our cohort of patients with ARDS, ARFA and CHF into two temporally independent datasets: a derivation dataset of 334 patients enrolled up to February 2019 and a validation dataset including 36 patients enrolled from March-August 2019. In secondary analyses, we also considered the 110 patients with acute on chronic repiratory failure, intubation for airway protection and

“other” category. Data for 235/334 (70%) of the patients in the derivation dataset had been previously utilized for application of LCA models separately in patients with ARDS and in patients ARFA (8).

We performed subphenotypic classifications by applying LCA in the derivation dataset. First, we estimated the optimal number of classes that best fit our patient cohort, as subphenotyping analysis has not yet been performed in such an inclusive patient population with acute respiratory failure. We considered a total of 35 baseline clinical and biomarker variables similar to those in LCA models previously used in ARDS subphenotyping trials (Additional File 2), without consideration of the clinical outcomes. The continuous variables were graphically examined by plotting their standardized values to a common z-scale (mean of 0, standard deviation of 1). We selected only variables that were found to be discriminatory ($p \geq 0.1$) for consideration in development of LCA and parsimonious models. Categorical variables were also compared graphically via Fisher exact tests with non-discriminatory variables being removed. We also applied the LCA to ARDS, ARFA, and CHF groups separately to confirm best fitness of the model in each individual subgroup.

We then developed a parsimonious logistic regression model based on a best subsets generalized linear model approach using Bayesian Information Criteria (BIC) (30). We subsequently applied subphenotype classifications provided from the parsimonious model in a) the derivation cohort and b) validation cohort of patients with ARDS, ARFA and CHF, and c) in the secondary analysis of patients with other forms of respiratory failure. Comparisons between hyper-inflammatory and hypo-inflammatory subphenotypes were obtained from Wilcoxon test for continuous variables and Fisher’s test for categorical variables. Kaplan-Meier curves and Cox-proportional hazard models were created for survival and time-to-liberation. We examined p-values for bootstrapped parametric likelihood ratio tests to select the final number of classes. We tested the proportional hazard assumption in all models. We performed LCA in Mplus 8.3 and all other analyses in R v.3.5.1 (31, 32).

Results:

Derivation Cohort Description:

We enrolled 334 patients (123 [37%] ARDS, 177 [53%] at-risk-for ARDS, 34 [10%] congestive heart failure) (Fig. 1). ARDS patients had the highest frequency of pneumonia, higher peak inspiratory pressures, and worse hypoxemia, whereas ARFA patients had higher incidence of aspiration and extra-pulmonary sepsis compared to the other groups ($p < 0.01$) (Additional File 3).

Subphenotypic classifications and baseline variables:

Utilizing all clinical and biomarker variables (n = 35), a two-class LCA model offered optimal fit ($p < 0.001$). Analysis of the LCA after exclusion of non-discriminatory variables (n = 11) improved the likelihood ratio of the model and also demonstrated excellent agreement with the model prior to elimination of these variables (Fig. 2). Twenty-nine percent of patients were assigned to the hyper-

inflammatory subphenotype, which had higher leukocytosis and creatinine levels, lower serum bicarbonate, and higher levels of all 10 measured biomarkers ($p < 0.01$) (Table 1).

Table 1
Comparisons of baseline variables between hyper-inflammatory and hypo-inflammatory patients by latent class analysis model.

Variable	Hypo-inflammatory	Hyper-inflammatory	p-value
	n = 236	**n = 98**	
Demographics			
Age (median [IQR])	59.4 [47.6, 68.1]	57.8 [45.5, 66.2]	0.75
Male gender, N (%)	131 (55.5)	57 (58.2)	0.2
BMI (median [IQR])	29.7 [24.9, 36.4]	28.3 [24.8, 34.0]	0.17
History of Chronic Disease			
Diabetes, N (%)	88 (37.3)	31 (31.6)	0.39
COPD, N (%)	59 (25.0)	20 (20.4)	0.45
Immunosuppression, N (%)	51 (21.6)	21 (21.4)	1
Chronic kidney disease, N (%)	31 (13.1)	28 (28.6)	**< 0.01**
Chronic cardiac failure, N (%)	31 (13.1)	11 (11.2)	0.77
Alcohol use, N (%)	33 (14.0)	20 (20.4)	0.19
Risk Factors for ARDS			
Pneumonia, N (%)	122 (51.7)	40 (40.8)	0.09
Aspiration, N (%)	53 (22.5)	17 (17.3)	0.37
sepsis, N (%)	54 (22.9)	40 (40.8)	**< 0.01**
LIPS score (median [IQR])	5.5 [4.5, 7.0]	6.5 [5.5, 8.0]	**< 0.01**
Hemodynamic parameters			
HR (median [IQR])	88.5 [76.0, 103.0]	100.0 [85.2, 111.8]	**< 0.01**

P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotypes were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease, LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight.

Variable	Hypo-inflammatory	Hyper-inflammatory	p-value
SBP (median [IQR])	117.0 [103.8, 133.0]	108.0 [96.0, 123.8]	**< 0.01**
Laboratory parameters			
pHa (median [IQR])	7.4 [7.3, 7.4]	7.3 [7.3, 7.4]	**< 0.01**
WBC (median [IQR])	12.4 [8.5, 16.9]	16.1 [10.9, 23.9]	**< 0.01**
Creatinine (median [IQR])	1.1 [0.7, 1.7]	3.2 [2.1, 4.5]	**< 0.01**
Serum CO2 (median [IQR])	24.0 [21.8, 28.0]	20.0 [18.0, 23.0]	**< 0.01**
Mechanical Ventilation Parameters			
Worst PaO2:FiO2 ratio (median [IQR])	163.0 [108.0, 206.5]	168.5 [118.5, 231.2]	0.38
Peak Inspiratory pressure (median [IQR])	25.0 [20.0, 31.2]	26.0 [21.0, 32.0]	0.43
Tidal Volume (per kg of PBW), median [IQR], ml/kg	6.7 [6.0, 7.5]	6.7 [6.0, 7.4]	0.54
Oxygen Saturation, median [IQR], %	97.0 [95.0, 99.0]	97.0 [95.0, 98.8]	0.47
Biomarkers			
IL-6, median [IQR], pg/ml	57 [23, 160]	235 [58, 711]	**< 0.01**
IL-8, median [IQR], pg/ml	18 [9, 29]	40 [24, 85]	**< 0.01**
IL-10, median [IQR], pg/ml	1 [1, 4]	11 [1, 23]	**< 0.01**
TNFR-1, median [IQR], pg/ml	3547 [2038, 5383]	10586 [7570, 16860]	**< 0.01**
Angiopoietin-2, median [IQR], pg/ml	6680 [3786, 12098]	20947 [15386, 36477]	**< 0.01**

P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotypes were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease, LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight.

Variable	Hypo-inflammatory	Hyper-inflammatory	p-value
Pentraxin-3, median [IQR], pg/ml	3471 [1201, 9348]	7586 [3656, 20103]	**< 0.01**
Fractalkine, median [IQR], pg/ml	1514 [908, 2306]	2808 [2124, 4038]	**< 0.01**
ST-2, median [IQR], pg/ml	159204 [74176, 377053]	622752 [233588, 1351038]	**< 0.01**
Procalcitonin, median [IQR], pg/ml	706 [266, 2045]	4353 [2009, 4966]	**< 0.01**
Rage, median [IQR], pg/ml	3122.4 [1783.8, 5218.1]	7846 [5359, 12814]	**< 0.01**
BDG [IQR], pg/ml	22 [14, 38]	42 [22, 73]	**< 0.01**

P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotypes were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease, LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight.

Proportions of the hyper- and hypo-inflammatory subphenotypes remained similar within each clinical subgroup (32% vs. 68% in ARDS, 28%. vs. 72% in ARFA, and 27% vs. 73% in CHF; p = 0.75) (Fig. 3). To confirm a two-class model offered best fit within each clinical group, the LCA was applied to the ARDS, ARFA, and CHF subgroups individually and was confirmed to offer optimal fit. Thus, LCA models confirmed the presence of hyper- and hypo-inflammatory subphenotypes in ARDS and ARFA groups and established subphenotypic presence in patients with acute respiratory failure secondary to CHF.

In derivation of the parsimonious model, 11 variables were found to be the most discriminatory between the two subphenotypes (Fig. 2): hypoalbuminemia, creatinine, bicarbonate, arterial pH, procalcitonin, RAGE, TNFR-1, ST-2, angiopoietin-2, fractalkine and pentraxin-3. With feature selection using a best subsets generalized linear model, we derived a 4-variable parsimonious model consisting of angiopoietin-2, TNFR-1, procalcitonin and bicarbonate, which offered excellent classification (area under the curve [AUC] = 0.98) against LCA-defined subphenotypes. This model predicted the probability of assignment to the hypo-inflammatory subphenotype: *0.8739604–8.798345e-05*(angiopoietin-2) – 6.049412e-04* (procalcitonin) – 4.048723e04(TNFR-1) + 2.883218e-01*(bicarbonate)*. The threshold probability for classification in to the hypo-inflammatory subphenotype was > 50%.

Analysis of the parsimonious model predictions against LCA subphenotypes in each individual clinical group offered excellent classification (AUC = 0.98 [0.93–0.99] for ARDS, 0.97 [0.96–0.99] for ARFA, and 0.95 [0.76-1.00] for CHF). Proportion of hyper- and hypo-inflammatory phenotypes and comparison of

baseline clinical variables and biomarkers between subphenotypes were very similar to those of the LCA models (Fig. 4, Additional file 4).

Secondary analysis in the remaining combined clinical groups (acute on chronic respiratory failure, airway protection, and “other”; n = 110) by the parsimonious predictive model demonstrated a very low prevalence of the hyper-inflammatory subphenotype (n = 8, 7%), which was significantly lower compared to the ARDS, ARFA and CHF groups combined ($p < 0.01$; Fig. 3).

Severity of illness and clinical outcomes by subphenotypes:

For LCA-derived subphenotypes, hyper-inflammatory patients had higher SOFA scores, vasopressor usage, AKI, 30- and 90-day mortality and fewer VFDs ($p < 0.01$) (Table 2). Similar associations with severity of illness and outcomes were observed for the parsimonious model predicted subphenotypes. Hyperinflammatory patients had worse 30-day survival and longer times to liberation from mechanical ventilation for both LCA and parsimonious model classifications (Fig. 5).

Table 2
Severity of illness and clinical outcome comparisons between LCA-derived subphenotypes and parsimonious derived subphenotypes.

LCA-derived			
Variable	Hypo-inflammatory	Hyper-inflammatory	p-value
N	236	98	
Shock (vasopressor use), N (%)	122 (51.7)	73 (74.5)	**< 0.01**
SOFA score (median [IQR])	7.0 [4.0, 9.0]	10.0 [7.2, 12.0]	**< 0.01**
Acute kidney injury, N (%)	149 (63.1)	90 (91.8)	**< 0.01**
30-d mortality, N (%)	57 (24.2)	40 (40.8)	**< 0.01**
90-d mortality, N (%)	65 (27.5)	43 (43.9)	**0.01**
ICU length of stay (median [IQR])	9.0 [5.0, 14.0]	10.0 [6.2, 14.0]	0.31
Ventilator-free days (median [IQR])	19.0 [0.0, 24.0]	1.0 [0.0, 21.0]	**< 0.01**
Duration of mechanical ventilation, median [IQR], d	6.0 [4.0, 12.0]	7.0 [4.0, 11.0]	0.42
Parsimonious-derived			
Variable	Hypo-inflammatory	Hyper-inflammatory	p-value
N	238	96	
Shock (vasopressor use), N (%)	123 (51.7)	73 (75.0)	**< 0.01**
SOFA score (median [IQR])	7.0 [4.0, 9.0]	10.0 [7.0, 12.0]	**< 0.01**
Acute kidney injury, N (%)	149 (62.6)	90 (93.8)	**< 0.01**
30-d mortality, N (%)	56 (23.5)	41 (42.7)	**< 0.01**
90-d mortality, N (%)	63 (26.5)	45 (46.9)	**< 0.01**
ICU length of stay (median [IQR])	9.0 [5.0, 15.0]	8.5 [5.8, 14.0]	1
Ventilator-free days (median [IQR])	19.0 [0.0, 24.0]	3.5 [0.0, 22.0]	**< 0.01**
Duration of mechanical ventilation, median [IQR], d	7.0 [4.0, 12.0]	6.0 [4.0, 10.8]	> 0.9

P-values for comparisons were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold.
Abbreviations: IQR: interquartile range; SOFA: sequential organ failure assessment.

LCA-derived
Abbreviations: BMI: body mass index; COPD: chronic obstructive pulmonary disease, PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; SBP: systolic blood pressure; WBC: white blood cell count; PEEP: positive end-expiratory pressure; PBW: predicted body weight
Additional File 3. Baseline characteristics and clinical outcomes by clinical diagnosis category (ARDS, ARFA, CHF). (DOCX)
P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotype were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease, LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; SOFA: sequential organ failure assessment; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight.
Additional File 4. Comparisons of baseline variables between hyper-inflammatory and hypo-inflammatory patients by the 4-variable parsimonious predictive model. (DOCX)
P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotype were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease, LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight.
Additional File 5: Validation cohort comparisons of baseline variables and outcomes by the 4-variable parsimonious predictive model.
P-values for comparisons between hyper-inflammatory and hypo-inflammatory subphenotype were obtained from Wilcoxon test for continuous variables and Fisher's test for categorical variables. Statistically significant p-values ($p < 0.05$) are highlighted in bold. Abbreviations: IQR: interquartile range; BMI: body mass index; COPD: chronic obstructive pulmonary disease; LIPS: lung injury prediction score; HR: heart rate; SBP: systolic blood pressure; WBC: white blood cell count; PaO_2: partial pressure of arterial oxygen; FiO_2: Fractional inhaled concentration of oxygen; PBW: predicted body weight; SOFA: sequential organ failure assessment; ICU: intensive care unit.

Validation cohort:

In a small validation cohort of 36 patients, we found a similar prevalence of the hyper-inflammatory subphenotype (22%) compared to the derivation cohort. Hyperinflammatory patients in the validation cohort had numerically higher 30-day mortality (25% vs. 11%) and statistically significantly higher levels of biomarkers not included in the parsimonious model (e.g. RAGE, IL-6 and ST-2) compared to hypoinflammatory patients (Additional File 5).

Discussion:

In a prospective, observational cohort of mechanically-ventilated patients with acute respiratory failure, we demonstrated that the subphenotypic distinction described for patients with ARDS applies in a broader and more heterogeneous critically-ill population (5, 33). While previous subphenotyping models have mainly been derived from secondary analyses of clinical trial populations in patients with ARDS, we revealed the presence of such subphenotypes in an observational, inclusive cohort study. We generated a simplified LCA model utilizing only discriminatory variables at baseline, and then defined a parsimonious 4-variable regression model consisting of well-described biomarkers in ARDS and other critical illness syndromes. Consistent with prior studies, we found that hyper-inflammatory patients at baseline exhibited higher severity of illness, elevated plasma biomarkers beyond the ones used for the subphenotypic predictions, and worse clinical outcomes. Our model encompasses a broad range of patients presenting with acute respiratory failure and various clinical manifestations, findings that support the use of biomarker-based subphenotyping in broader critically ill patient populations.

Subphenotyping research has gained wide popularity in recent years, not only within ARDS but within other critical care syndromes such as sepsis and AKI. Amongst these syndromes, attempts to subdivide patients solely based on clinical criteria have been largely unsuccessful when compared to biomarker-based modeling. In ARDS, subgroups receiving differential ventilator management strategies based on CT morphology showed no difference in mortality, and this strategy even proved to be harmful if misclassified into the wrong group (34). In AKI, similar to ARDS, no effective therapeutic interventions exist other than supportive care and renal replacement therapy. However, with biomarker-based subphenotyping, two distinct subphenotypes of AKI emerged with differential responses to therapy such as vasopressin (35). Parsimonious modeling revealed that the distinguishing biomarkers in AKI are extremely similar to those of our cohort, including bicarbonate, angiopoietin-2, and TNFR1, findings which raise the possibility of common pathways of systemic inflammatory responses between heterogeneous populations of critically-ill patients.

An important challenge in critical care pertains to the subjective and non-specific nature of diagnostic criteria for critical illness syndromes. ARDS is systematically underrecognized or under-reported as a diagnosis in clinical practice; and even among expert providers, there are significant rates of diagnostic disagreement (11, 36). ARDS recognition is straightforward in cases with typical presentations of diffuse, bilateral infiltrates on imaging with an obvious risk factor, such as pneumonia or sepsis. However less classic radiographic presentations are a source of uncertainty and diagnostic discordance. In a retrospective analysis where expert panels identified ARDS by electronic medical record review, ARDS diagnosis had been documented in the clinical chart only 12.4% of the time, at least in part due to underrecognition of the syndrome at the time of clinical encounter (12). Further investigation into underrecognition of ARDS demonstrated that interobserver agreement of ARDS diagnosis under Berlin criteria has only been moderate, with lack of consensus on chest radiograph interpretation accounting for most differences (11). Inter-rate agreement in ARDS radiographic identification remained low even after educational interventions (9). Given such inherent diagnostic uncertainty in ARDS, there is compelling need to understand the underlying biological mechanisms that drive different outcomes in these

conditions. Our parsimonious model is not dependent on the requirement of a definite clinical diagnosis (i.e. ARDS vs. CHF) and allows for objective classification into two distinctly behaving subphenotypes.

Our LCA and parsimonious models were created after combining ARDS, ARFA, and CHF patients. While the sample sizes of each group within our study differ, we confirmed that the same hyper- and hypo-inflammatory subphenotypes exist in all 3 subgroups independently, mitigating concern that any one group may be the dominating factor driving subphenotypic results. The remaining respiratory failure groups demonstrated a very low prevalence of hyper-inflammatory patients (7%), confirming that we included only clinically pertinent groups in our primary analysis. Patients intubated for airway protection or those with acute exacerbation for COPD appear to have extremely low prevalence of a hyper-inflammatory subphenotype, and thus this classification framework would not be relevant in these forms of acute respiratory failure.

CHF is a clinical group with paucity of data on subphenotyping, although it has already been proposed that treatment of systolic heart failure should be shifted away from strategies that simply improve cardiac function towards interventions that modulate the systemic responses to cardiac dysfunction instead (37). Interestingly, our results demonstrate that hyper- and hypo-inflammatory subphenotypes exist in CHF patients in very similar proportion to ARDS and ARFA patients, further suggesting biological commonalities amongst critical illness syndromes, regardless of clinical diagnosis. Exclusion of cardiac edema is a major source of disagreement in the clinical decision to diagnose ARDS, and as seen in the FACTT trial, a significant amount of patients diagnosed with ARDS have elevated pulmonary capillary wedge pressures (38). Consequently, it is probable that a certain unknown proportion of patients with ARDS are misclassified as cardiogenic pulmonary edema, and vice versa. A recent study evaluated the longitudinal evolution of radiographic pulmonary edema in ARDS and found that baseline radiographic pulmonary edema was not associated with ARDS severity, clinical outcomes or hyper-inflammatory subphenotype (13). These findings further highlight the disadvantage in using clinical criteria rather than objective biological markers to subdivide such populations.

Our 4-variable parsimonious model includes two markers that have been associated in cohorts much larger than our current study (bicarbonate and TNFR1) (8, 20), one well-validated marker in sepsis and ARDS (angiopoeitin-2) (23, 24), and one already widely clinically available and utilized test (procalcitonin) (27, 28). Angiopoietin-2 may be a possible causal factor in development of ARDS in septic patients (24), in addition to an independent predictor of increased mortality in ARDS (25, 39). Angiopoietin-2 has been included in prior parsimonious models (25, 35, 40) for ARDS and AKI subphenotypes. Similarly, bicarbonate and TNFR1 have been identified as key predictors in prior parsimonious models for ARDS (6, 41). Notably, these analyses did not include the other two biomarkers in our parsimonious model (angiopoietin-2, procalcitonin) as possible classifier variables. Further verification in larger data sets is therefore required. While a parsimonious model allows for easier clinical applicability than complex LCA, models that depend on biomarker variables will eventually require point-of-care or rapid turnaround tests to ensure timely acquisition of results and application to model predictions.

Our study utilized a simplified approach to variable selection for the LCA models. Prior studies have utilized large amounts of variables without any selectivity. In our approach, we removed non-discriminatory variables prior to analysis and found excellent agreement between models before and after removal (Fig. 2). These findings demonstrate the ability to perform LCA with a smaller set of variables and similar or improved statistical performance of the model.

Our study has several limitations. While our study prospectively enrolled consecutive patients with acute respiratory failure, it is limited by sample size and single center design. We also only examined baseline data (within 48 hours intubation), therefore it is unclear whether patients transition between subphenotypes over time. In one analysis, from day 0 to day 3 the majority (> 94%) of patients remained in the same subphenotype (42), though preliminary data within our own cohort demonstrate the possibility of higher rates of transition by days 3–6 (43). Further examination of the stability of subphenotypes will be important, as transition from one group to another will affect the ability to target clinical interventions. While we had an independent data set for validation demonstrating similar trends with outcomes biomarkers, this validation dataset is small, illustrating the need to validate our model in larger data sets. Additionally, we have not yet investigated any differential therapeutic responses between groups, which will be required for clinical applications. Larger, prospective trials of heterogeneous critically ill populations will be necessary to verify subphenotypic models and and explore treatment effects.

Conclusions:

The formerly described ARDS hyper- and hypo-inflammatory subphenotypes can be observed in a broader and more heterogeneous population of patients with acute respiratory failure, offering prognostic enrichment and suggesting underlying biological commonalities amongst diverse clinical syndromes regardless of specific diagnosis. Objective and accurate subphenotyping of these patients is possible with the use of simple predictive models based on biomarker values.

Abbreviations

AIC (Akaike information criterion), AKI (acute kidney injury), Ang-2 (angiopoietin-2), ARDS (acute respiratory distress syndrome), ARFA (at-risk-for ARDS), ATS (American Thoracic Society), AUC (area under the curve), BDG (beta-D-glucan), BIC (Bayesian information criterion), CHF (congestive heart failure), COPD (chronic obstructive pulmonary disease), FACTT (Fluids and Catheters Treatment Trial), ICU (intensive care unit), IL (interleukin), ILD (interstitial lung disease), IQR (interquartile range), LCA (latent class analysis), RAGE (receptor of advanced glycation end-products), SOFA (sequential organ failure assessment), ST-2 (suppression of tumorigenicity-2), TNFR1 (soluble tumor necrosis factor receptor-1), VFD (ventilator free days)

Declarations

Ethics approval and consent to participate: The study was approved by the University of Pittsburgh Institutional Review Board (protocol STUDY19050099). Written informed consent was provided by all participants or their surrogates.

Consent for publication: Not applicable

Availability of data and materials: Most of the data used and analyzed during this study are included in this published article and its supplementary information files. The datasets used during the current study are available from the corresponding author upon reasonable request.

Competing interests: Dr. Bryan J. McVerry has been a consultant for Vapothern, Inc. and receives research funding from Bayer Pharmaceuticals, Inc. Dr. Georgios Kitsios has received funding from Karius, Inc. The other authors have no competing interests to declare.

Funding: National Institutes of Health [K23 HL139987 (GDK); P01 HL114453 (BJM); R01 HL097376 (BJM); U01 HL098962 (AM); K24 HL123342 (AM); F32 HL137258 (JE); F32 HL142172 (WB); K23GM122069 (FS)].

Author Contributions:

Conception and design: CMD, GDK, BJM, SMN

Acquisition, analysis or interpretation of data: WB, FS, JE, YZ, AM, CMD, SMN, GDK, BJM

Clinical cohort phenotyping: CMD, SMN, BJM, GDK

Drafting of work and/or revising for important intellectual content: CMD, SMN, BJM, GDK

Final approval of version to be published: CMD, SMN, WB, FAS, JE, YZ, AM, BJM, GDK

Acknowledgements: Not applicable

References

1. Rubenfeld GD. Confronting the frustrations of negative clinical trials in acute respiratory distress syndrome. Ann Am Thorac Soc. 2015;12(Suppl 1):58–63.
2. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, et al. Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. JAMA. 2016;315(8):788–800.
3. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of Heterogeneity of Treatment Effect for Reporting and Analysis of Randomized Trials in Critical Care. Am J Respir Crit Care Med. 2015;192(9):1045–51.

4. Prescott HC, Calfee CS, Thompson BT, Angus DC, Liu VX. Toward Smarter Lumping and Smarter Splitting: Rethinking Strategies for Sepsis and Acute Respiratory Distress Syndrome Clinical Trial Design. Am J Respir Crit Care Med. 2016;194(2):147–55.
5. Reddy K, Sinha P, O'Kane CM, Gordon AC, Calfee CS, McAuley DF. Subphenotypes in critical care: translation into clinical practice. Lancet Respir Med. 2020;8(6):631–43.
6. Sinha P, Delucchi KL, McAuley DF, O'Kane CM, Matthay MA, Calfee CS. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. Lancet Respir Med. 2020;8(3):247–57.
7. Sinha P, Churpek MM, Calfee CS. Machine Learning Classifier Models Can Identify ARDS Phenotypes Using Readily Available Clinical Data. Am J Respir Crit Care Med. 2020.
8. Kitsios GD, Yang L, Manatakis DV, Nouraie M, Evankovich J, Bain W, et al. Host-Response Subphenotypes Offer Prognostic Enrichment in Patients With or at Risk for Acute Respiratory Distress Syndrome. Crit Care Med. 2019;47(12):1724–34.
9. Goddard SL, Rubenfeld GD, Manoharan V, Dev SP, Laffey J, Bellani G, et al. The Randomized Educational Acute Respiratory Distress Syndrome Diagnosis Study: A Trial to Improve the Radiographic Diagnosis of Acute Respiratory Distress Syndrome. Crit Care Med. 2018;46(5):743–8.
10. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, et al. Acute respiratory distress syndrome: the Berlin Definition. JAMA. 2012;307(23):2526–33.
11. Sjoding MW, Hofer TP, Co I, Courey A, Cooke CR, Iwashyna TJ. Interobserver Reliability of the Berlin ARDS Definition and Strategies to Improve the Reliability of ARDS Diagnosis. Chest. 2018;153(2):361–7.
12. Weiss CH, Baker DW, Weiner S, Bechel M, Ragland M, Rademaker A, et al. Low Tidal Volume Ventilation Use in Acute Respiratory Distress Syndrome. Crit Care Med. 2016;44(8):1515–22.
13. Kotok D, Yang L, Evankovich JW, Bain W, Dunlap DG, Shah F, et al. The evolution of radiographic edema in ARDS and its association with clinical outcomes: A prospective cohort study in adult patients. J Crit Care. 2020;56:222–8.
14. Kitsios GD, Yang H, Yang L, Qin S, Fitch A, Wang XH, et al. Respiratory Tract Dysbiosis is Associated With Worse Outcomes in Mechanically-Ventilated Patients. Am J Respir Crit Care Med. 2020.
15. Bain W, Li H, van der Geest R, Moore SR, Olonisakin TF, Ahn B, et al. Increased Alternative Complement Pathway Function and Improved Survival during Critical Illness. Am J Respir Crit Care Med. 2020;202(2):230–40.
16. McKay HS, Margolick JB, Martínez-Maza O, Lopez J, Phair J, Rappocciolo G, et al. Multiplex assay reliability and long-term intra-individual variation of serologic inflammatory biomarkers. Cytokine. 2017;90:185–92.
17. Liu CH, Kuo SW, Ko WJ, Tsai PR, Wu SW, Lai CH, et al. Early measurement of IL-10 predicts the outcomes of patients with acute respiratory distress syndrome receiving extracorporeal membrane oxygenation. Sci Rep. 2017;7(1):1021.

18. Bajwa EK, Volk JA, Christiani DC, Harris RS, Matthay MA, Thompson BT, et al. Prognostic and diagnostic value of plasma soluble suppression of tumorigenicity-2 concentrations in acute respiratory distress syndrome. Crit Care Med. 2013;41(11):2521–31.
19. Hoogendijk AJ, Wiewel MA, van Vught LA, Scicluna BP, Belkasim-Bohoudi H, Horn J, et al. Plasma fractalkine is a sustained marker of disease severity and outcome in sepsis patients. Crit Care. 2015;19:412.
20. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA, et al. Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. Lancet Respir Med. 2014;2(8):611–20.
21. Jabaudon M, Blondonnet R, Pereira B, Cartin-Ceba R, Lichtenstern C, Mauri T, et al. Plasma sRAGE is independently associated with increased mortality in ARDS: a meta-analysis of individual patient data. Intensive Care Med. 2018;44(9):1388–99.
22. Jones TK, Feng R, Kerchberger VE, Reilly JP, Anderson BJ, Shashaty MGS, et al. Plasma sRAGE Acts as a Genetically Regulated Causal Intermediate in Sepsis-associated Acute Respiratory Distress Syndrome. Am J Respir Crit Care Med. 2020;201(1):47–56.
23. Calfee CS, Gallagher D, Abbott J, Thompson BT, Matthay MA, Network NA. Plasma angiopoietin-2 in clinical acute lung injury: prognostic and pathogenetic significance. Crit Care Med. 2012;40(6):1731–7.
24. Reilly JP, Wang F, Jones TK, Palakshappa JA, Anderson BJ, Shashaty MGS, et al. Plasma angiopoietin-2 as a potential causal marker in sepsis-associated ARDS development: evidence from Mendelian randomization and mediation analysis. Intensive Care Med. 2018;44(11):1849–58.
25. Bime C, Casanova N, Oita RC, Ndukum J, Lynn H, Camp SM, et al. Development of a biomarker mortality risk model in acute respiratory distress syndrome. Crit Care. 2019;23(1):410.
26. Mauri T, Coppadoro A, Bellani G, Bombino M, Patroniti N, Peri G, et al. Pentraxin 3 in acute respiratory distress syndrome: an early marker of severity. Crit Care Med. 2008;36(8):2302–8.
27. Liu D, Su LX, Guan W, Xiao K, Xie LX. Prognostic value of procalcitonin in pneumonia: A systematic review and meta-analysis. Respirology. 2016;21(2):280–8.
28. Yu Z, Ji M, Hu X, Yan J, Jin Z. [Value of procalcitonin on predicting the severity and prognosis in patients with early ARDS: a prospective observation study]. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue. 2017;29(1):34–8.
29. Kitsios GD, Kotok D, Yang H, Finkelman M, Zhang Y, Britton N, et al. Plasma (1–3)-β-d-glucan levels are associated with host inflammatory responses and predict adverse clinical outcomes in critical illness. medRxiv2020.
30. Jovanovic BD, Hosmer DW, Buonaccorsi JP. Equivalence of several methods for efficient best subsets selection in generalized linear models. 1 ed. Computational Statistics & Data Analysis1995. p. 59–64.
31. R Foundation for Statistical Computing RCT. R: A Language and Environment for Statistical Computing. Vienna: CRAN; 2016.

32. Muthén LK, Muthén BO. Mplus User's Guide.. Los Angeles, CA: Muthén & Muthén; 1997–2017.
33. Sinha P, Delucchi KL, Thompson BT, McAuley DF, Matthay MA, Calfee CS, et al. Latent class analysis of ARDS subphenotypes: a secondary analysis of the statins for acutely injured lungs from sepsis (SAILS) study. Intensive Care Med. 2018;44(11):1859–69.
34. Constantin JM, Jabaudon M, Lefrant JY, Jaber S, Quenot JP, Langeron O, et al. Personalised mechanical ventilation tailored to lung morphology versus low positive end-expiratory pressure for patients with acute respiratory distress syndrome in France (the LIVE study): a multicentre, single-blind, randomised controlled trial. Lancet Respir Med. 2019;7(10):870–80.
35. Bhatraju PK, Zelnick LR, Herting J, Katz R, Mikacenic C, Kosamo S, et al. Identification of Acute Kidney Injury Subphenotypes with Differing Molecular Signatures and Responses to Vasopressin Therapy. Am J Respir Crit Care Med. 2019;199(7):863–72.
36. Sjoding MW, Hofer TP, Co I, McSparron JI, Iwashyna TJ. Differences between Patients in Whom Physicians Agree and Disagree about the Diagnosis of Acute Respiratory Distress Syndrome. Ann Am Thorac Soc. 2019;16(2):258–64.
37. Lawler PR, Mehra MR. Advancing from a "hemodynamic model" to a "mechanistic disease-modifying model" of cardiogenic shock. J Heart Lung Transplant. 2018;37(11):1285–8.
38. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, deBoisblanc B, et al. Comparison of two fluid-management strategies in acute lung injury. N Engl J Med. 2006;354(24):2564–75.
39. Terpstra ML, Aman J, van Nieuw Amerongen GP, Groeneveld AB. Plasma biomarkers for acute respiratory distress syndrome: a systematic review and meta-analysis*. Crit Care Med. 2014;42(3):691–700.
40. Bos LD, Schouten LR, van Vught LA, Wiewel MA, Ong DSY, Cremer O, et al. Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. Thorax. 2017;72(10):876–83.
41. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, et al. Acute Respiratory Distress Syndrome Subphenotypes Respond Differently to Randomized Fluid Management Strategy. Am J Respir Crit Care Med. 2017;195(3):331–8.
42. Delucchi K, Famous KR, Ware LB, Parsons PE, Thompson BT, Calfee CS, et al. Stability of ARDS subphenotypes over time in two randomised controlled trials. Thorax. 2018;73(5):439–45.
43. Drohan C, Nouraie SM, Shah F, Yang L, Bain W, Evankovich J, et al. Longitudinal Evolution of Host-Response Subphenotypes in Critical Illness. American Journal of Respiratory and Critical Care Medicine2020.

Figures

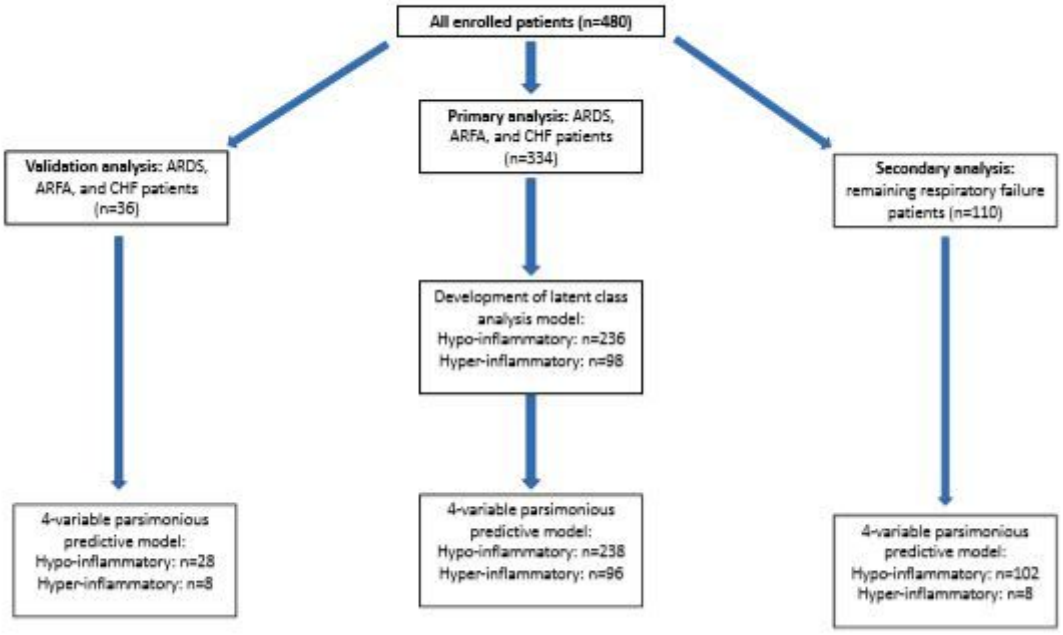

Figure 1

Flow chart of the study Flow chart depicts primary analysis in the center, compared to the validation cohort (left) and secondary analysis of other acute respiratory failure groups (right).

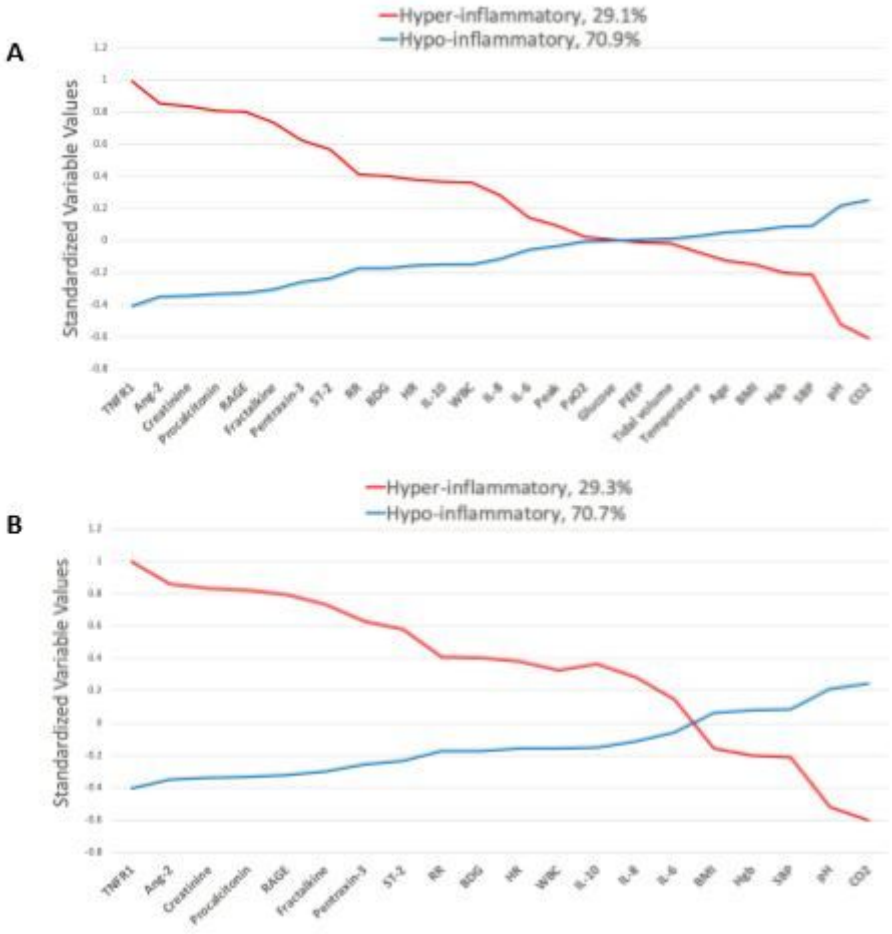

Figure 2

Differences of standardized values of continuous variables by LCA phenotype Differences of standardized values are displayed prior to removal (A) and after removal (B) of 7 non-discriminatory continuous variables. The variables are sorted from the highest degree of positive separation between subphenotypes on the left (hyper-inflammatory higher than hypo-inflammatory). The variables were standardized to a common z-scale (mean of 0, standard deviation of 1). Entropy for method A vs. B is similar (0.89 vs. 0.88; $p<0.001$). Categorical variables were also compared graphically via Fisher exact tests with 4 non-discriminatory variables being removed (not shown). Abbreviations: RR: respiratory rate; COPD: chronic obstructive pulmonary disease, PaO2: partial pressure of arterial oxygen; FiO2: Fractional inhaled concentration of oxygen; WBC: white blood cell count; Peak: Peak inspiratory pressure; PaO2: partial pressure of arterial oxygen; PEEP: positive end-expiratory pressure; BMI: body mass index; Hgb: hemoglobin; SBP: systolic blood pressure.

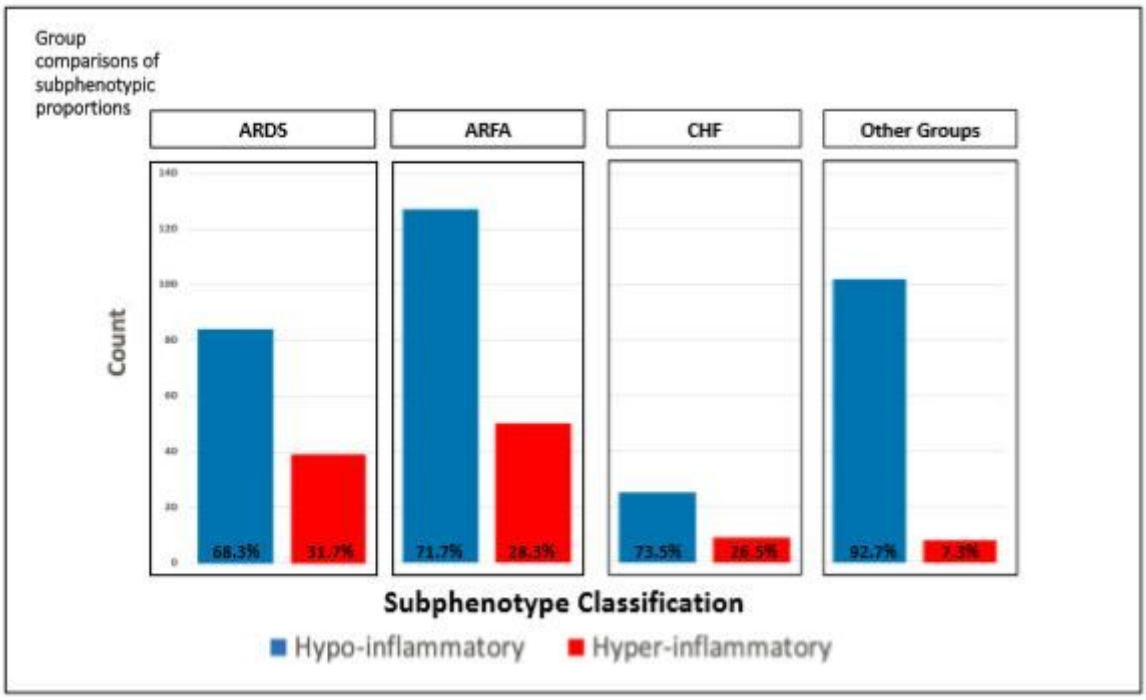

Figure 3

Distribution of subphenotypes among patients with ARDS, ARFA, CHF, and other groups. There was no significant difference in proportion of patient subphenotypes between ARDS, at-risk-for ARDS (ARFA), and CHF patients ($p=0.75$). A minority of patients classified within the other acute hypoxic respiratory groups were assigned to the hyper-inflammatory subphenotype (7%) ($p<0.01$). Between group comparisons of proportions were performed with a Fischer's exact test.

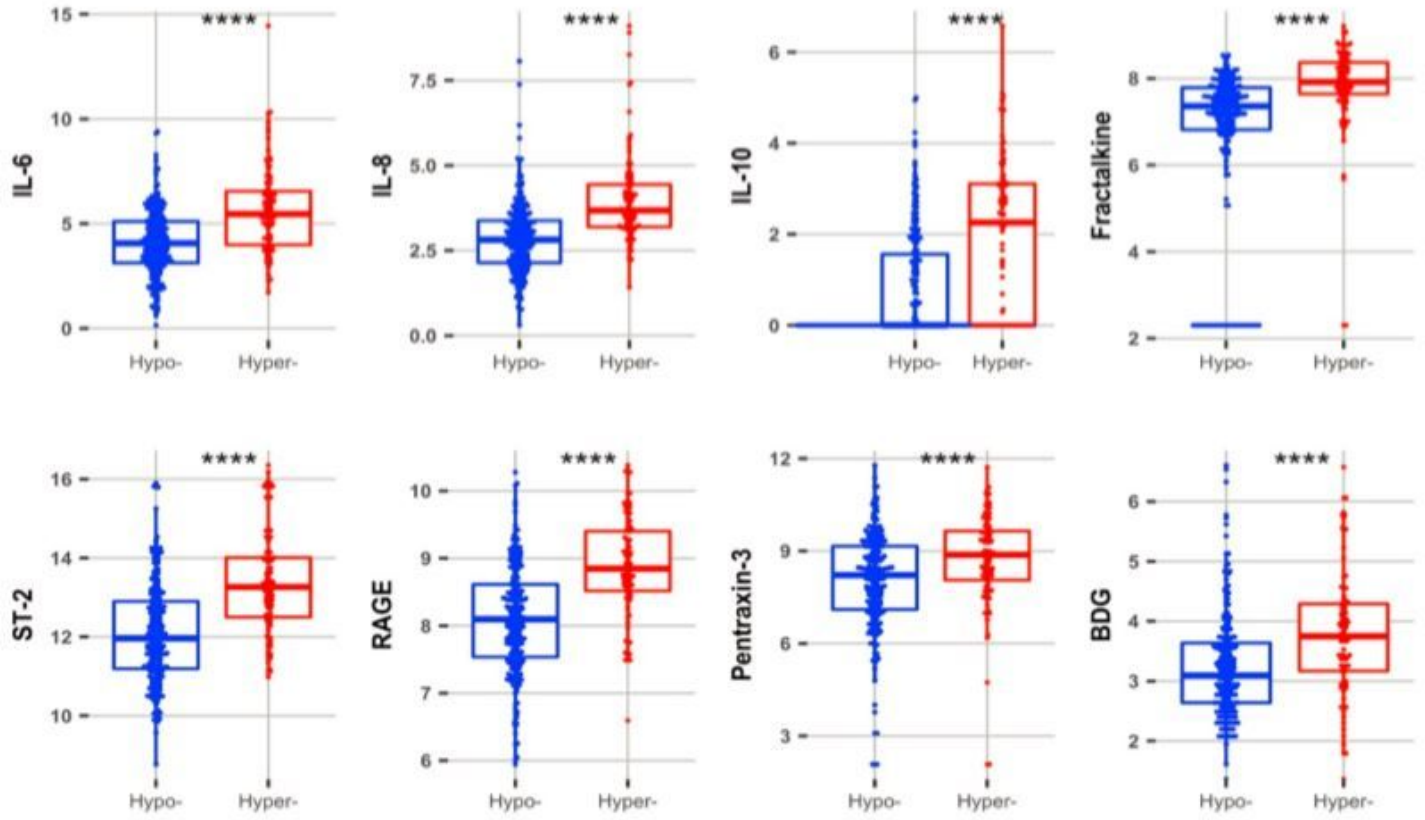

Figure 4

Box plot comparisons of biomarkers between subphenotypes Biomarkers not included in the prediction of the parsimonious model are shown as box plots comparing both subphenotypes. Biomarker data have been log-transformed. All biomarker comparisons were statistically significant between subphenotypes ($p<0.01$). Raw data and p-values are seen in Additional File 4.

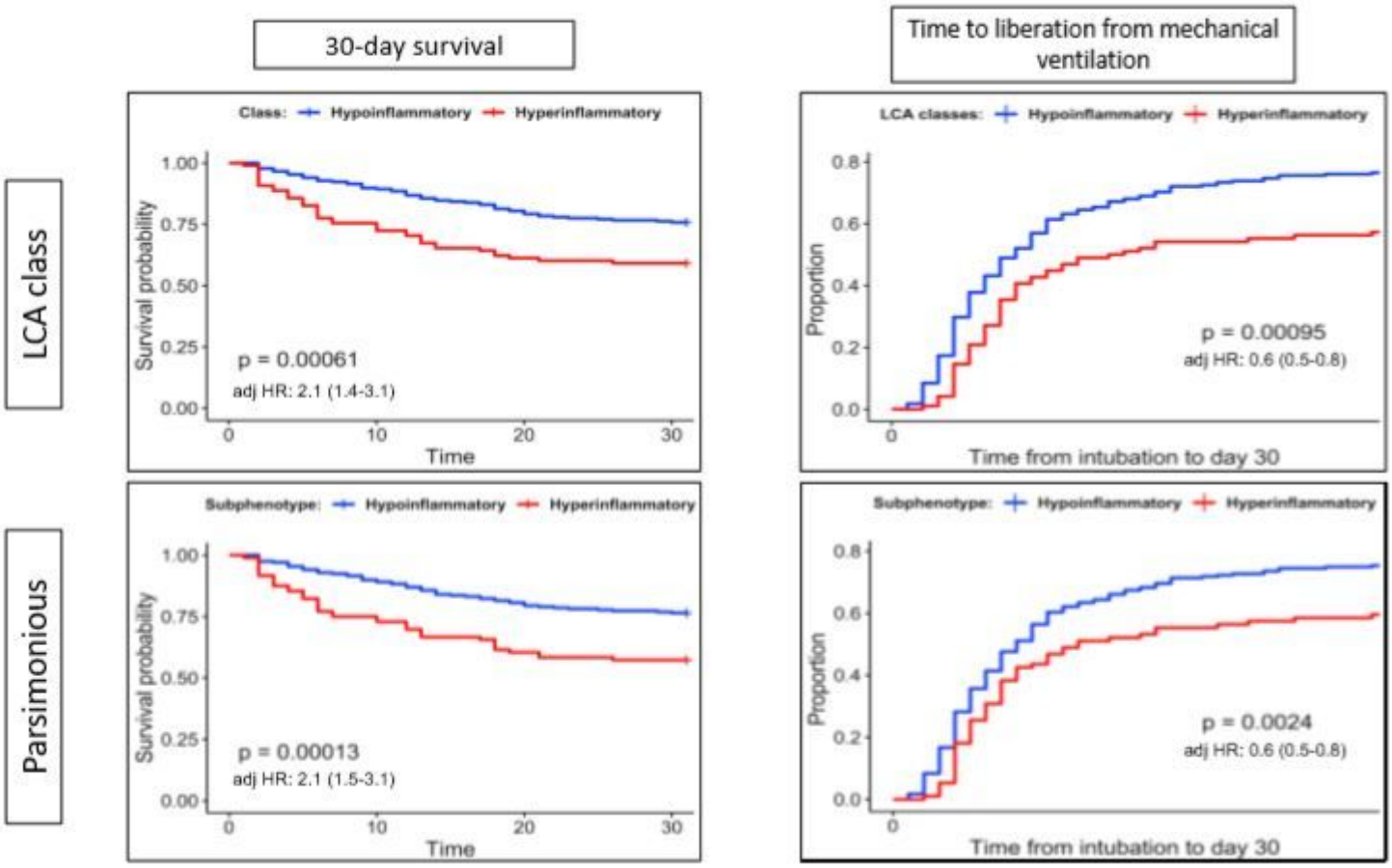

Figure 5

Hyper-inflammatory subjects have higher mortality and longer mechanical ventilation duration in LCA and parsimonious models. Kaplan Meier curves for 30-d survival (left panels) and time-to-liberation from mechanical ventilation (right panels) for each subphenotype as derived by LCA (top row) and the parsimonious 4-variable model (bottom row). P-values for differences between subphenotypes were obtained with a log-rank test. Adjusted hazard ratios (aHR) with 95% confidence intervals are displayed for the effects of the hyper-inflammatory subphenotype and were derived from multivariate Cox-proportional hazards models. All models were adjusted for age. 90-day survival data were very similar to 30-day and are not shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile4.docx
- AdditionalFile5.docx
- AdditionalFile2.docx

- AdditionalFile1.docx
- AdditionalFile3.docx