BMJ Open

Identification of Acute Respiratory Distress Syndrome subphenotypes denovo using routine clinical data: a retrospective analysis of ARDS clinical trials

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053297
Article Type:	Original research
Date Submitted by the Author:	11-May-2021
Complete List of Authors:	Duggal, Abhijit ; Cleveland Clinic, Department of Critical Care Medicine Kast, Rachel; Endpoint Health Inc, Department of Clinical Data Science Van Ark, Emily; Endpoint Health Inc, Department of Clinical Data Science Bulgarelli, Lucas; Endpoint Health Inc, Department of Clinical Data Science Siuba, Matthew T.; Cleveland Clinic, Department of Critical Care Medicine Osborn, Jeff; Endpoint Health Inc, Department of Clinical Data Science Rey, Diego; Endpoint Health Inc, Department of Clinical Data Science Zampieri, Fernando; Hospital do Coração Cavalcanti, Alexandre ; HCor Research Institute Maia, Israel; Hospital do Coracao Paisani, Denise M; HCor Research Institute Laranjeira, Ligia N; HCor Research Institute Serpa Neto, Ary; Monash University Deliberato, Rodrigo Octávio; Endpoint Health Inc, Department of Clinical Data Science
Keywords:	Adult intensive & critical care < INTENSIVE & CRITICAL CARE, Adult intensive & critical care < ANAESTHETICS, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Identification of Acute Respiratory Distress Syndrome subphenotypes denovo using routine clinical data: a retrospective analysis of ARDS clinical trials

Abhijit Duggal MD[1‡], Rachel Kast PhD[2‡], Emily Van Ark PhD[2‡], Lucas Bulgarelli BSc[2‡], Matthew T. Siuba DO[1], Jeff Osborn[2], Diego Rey PhD[2], Fernando G Zampieri MD PhD[3], Alexandre B Cavalcanti MD PhD[3], Israel S Maia MD[3], Denise M Paisani PhD[3], Ligia N Laranjeira[3], Ary Serpa Neto MD MSc PhD,[4,5,6,7,8] Rodrigo Octávio Deliberato MD PhD[2]

‡ Authors contributed equally

1. Department of Critical Care Medicine, Respiratory Institute, Cleveland Clinic, Cleveland, Ohio, USA.
2. Department of Clinical Data Science, Endpoint Health Inc, Palo Alto, California, USA.
3. HCor Research Institute, São Paulo, Brazil
4. Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.
5. Department of Critical Care, Melbourne Medical School, University of Melbourne, Austin Hospital, Melbourne, Australia.
6. Department of Intensive Care, Austin Hospital, Melbourne, Australia.
7. Data Analytics Research and Evaluation (DARE) Centre, Austin Hospital, Melbourne, Australia.
8. Department of Critical Care Medicine, Hospital Israelita Albert Einstein, São Paulo, Brazil.

Correspondence:
Abhijit Duggal MD
Address: 9500 Euclid Ave, L2-330, Cleveland, Ohio, 44195
E-mail: duggala2@ccf.org

Word count (Abstract): 235 words

Word count (Text): 2814 words

Number of figures: 2 figures

Number of tables: 3 tables

Supplementary Material: 01

Key words: Subphenotype, machine learning, ARDS, critical care, clinical data, clustering

ABSTRACT

Objectives: The acute respiratory distress syndrome (ARDS) is a heterogenous condition, and identification of subphenotypes may help in better risk stratification. Identify ARDS subphenotypes using new simpler methodology and readily available clinical variables using a retrospective analysis of previously published ARDS trials.

Setting: Data from the U.S. ARDSNet trials and from the international ART trial.

Participants: 3763 patients from ARDSNet datasets and 1010 patients from the ART dataset.

Primary and secondary outcome measures: The primary outcome was 60-day or 28-day mortality, depending on what was reported in the original trial. K-means cluster analysis was performed to identify subgroups. For feature selection, sets. Sets of candidate variables were tested to assess their ability to produce different probabilities for mortality in each cluster. Clusters were compared to biomarker data, allowing identification of subphenotypes.

Results: Data from 4,773 patients was analyzed. Two subphenotypes (A and B) resulted in optimal separation in the final model, which included nine routinely collected clinical variables, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. Participants in subphenotype B showed increased levels of pro-inflammatory markers, had consistently higher mortality, lower number of ventilator-free days at day 28, and longer duration of ventilation compared to patients in the subphenotype A.

Conclusions: Routinely available clinical data can successfully identify two distinct subphenotypes in adult ARDS patients. This work may facilitate implementation of precision therapy in ARDS clinical trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Largest cohort of patients used to identify subphenotypes of ARDS patients.
- Subphenotypes were validated in the population of a large international ARDS randomized controlled trial.
- Subphenotypes were identified by using only routinely collected clinical data.
- Our use of data exclusively from randomized controlled trials does not prove generalizability to unselected ARDS populations.
- The clinical utility of the subphenotypes have to be validated in a prospective study.

INTRODUCTION

The Berlin definition of acute respiratory distress syndrome (ARDS) encompasses acute hypoxemic respiratory failure due to a wide variety of etiologies [1]. Due to this inclusion of heterogeneous conditions within the syndrome, there are significant clinical and biological differences that makes ARDS challenging to treat [2,3]. These differences amongst ARDS patients are associated with variation in risk of disease development and progression [3,4], potentially generating differential responses to treatments and interventions [5–10]. In spite of those evidences, clinical risk stratification of ARDS patients still solely depends on PaO_2/FiO_2 ratios [11,12], possibly misleading the interpretation of results in clinical trials and clinicians when evaluating treatment options for patients [13].

Therefore, identifying groups of patients who have similar clinical, physiologic, or biomarker traits becomes relevant [6,14] as it can help with stratification of patients producing better targeted therapies and interventions [15]. These different groups can be defined as ARDS subphenotypes [4,14]. Two ARDS subphenotypes have been consistently identified in previous studies [6–10,16–18]. However, these models are complex, and significant barriers exist in their implementation and use in clinical practice. Existing models use up to 40 predictor variables, including biomarkers and other variables that are not readily available at the bedside [6–10,16–18]. These limitations explain the current status quo of ARDS care, where clinicians must depend on the limited prognostic value of PaO_2/FiO_2 ratios instead of biologically distinct subphenotypes.

We hypothesized that the use of a simpler methodology and a small number of easily available clinical variables could identify new ARDS subphenotypes and thus provide the means to allow future implementation of bedside stratification.

METHODS

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Data source and participants

We performed a retrospective study using a de-identified dataset pooling data from six randomized clinical trials in patients with ARDS, namely: ARMA, ALVEOLI, FACTT, EDEN, SAILS, and ART [19–24]. Patients in ARMA, ALVEOLI, FACTT, EDEN and SAILS trials were eligible if they met the American-European consensus for ARDS, including patients with a PaO_2 / FiO_2 ratio < 300 up to 48 hours before enrollment. From 1996 to 2013, these trials enrolled 902, 549, 1000, 1000, and 745 patients, respectively, and tested a variety of interventions [19–23]. Between 2011 and 2017 the international ART study enrolled 1010 adult patients diagnosed with moderate to severe ARDS according to the Berlin definition (PaO_2 / FiO_2 ratio < 200) for less than 72 hours of duration and assessed two different ventilatory strategies [24]. To avoid biases due to high mortality in the high tidal volume group of the ARMA study [19], which has not been standard of care since the beginning of 2000, only 473 patients receiving low tidal volume in that study were included.

Predictors

Six clinical trials were assessed to identify a set of clinical variables recorded closest to time of randomization which were most commonly available across all datasets. The list of potential candidates was then further refined to include only those that are frequently observed in the routine care of ARDS patients at the time of its diagnosis. In order to

develop a clustering algorithm for potential rapid translation into clinical use, elements which would not be commonly found in the electronic health records (EHR) at the time of ARDS diagnosis, such as biomarker levels, ARDS risk factors, organ support apart from mechanical ventilation settings, and severity scores, were excluded from model development. The treatment assignment in the original trials, and clinical outcomes were not considered in the model development.

After all assessment, 16 variables that are routinely collected as part of the usual care and which were uniformly present in all the trials were considered, including: age, gender, arterial pH, PaO_2, $PaCO_2$, bicarbonate, creatinine, bilirubin, platelets, heart rate, respiratory rate, mean arterial pressure, positive end-expiratory pressure (PEEP), plateau pressure, FiO_2, and tidal volume adjusted for predicted body weight (mL/kg PBW). The PBW was calculated as equal to 50 + 0.91 (centimeters of height – 152.4) in males, and 45.5 + 0.91 (centimeters of height – 152.4) in females [18]. These variables were grouped into five domains named demographics, arterial blood gases, laboratory values, vital signs, and ventilatory variables. Plateau pressure was excluded due to a high rate of missingness across the trials included in the training set. Amount of missing data in the training datasets is reported in **eTable 1**.

Outcomes

The primary outcome was 60-day mortality for all ARDSnet trials, and 28-day mortality for ART trial. Secondary outcomes included 90-day mortality, number of ventilator free days at day 28 [25], and the duration of mechanical ventilation in survivors within the first 28 days post enrollment.

Data preparation

Data preprocessing was performed before modeling, and the pooled dataset was assessed for completeness and consistency. Patients with values out of the plausible physiological range for a specific variable were excluded from the final analysis (described in **eTable 2**). The training dataset was constructed using data from the two largest ARDSnet trials, EDEN and FACTT. The validation dataset was sourced from the four remaining trials: ALVEOLI, ARMA, SAILS, and ART. Means and standard deviations for z-scoring variables were calculated from the training dataset and subsequently applied to the validation data.

Statistical analysis

Baseline and outcome data were presented according to the assigned cluster. Continuous variables were presented as medians with their interquartile ranges and categorical variables as total number and percentage. Proportions were compared using Fisher exact tests and continuous variables were compared using the Wilcoxon rank-sum test. Study outcomes were further compared using the median and mean absolute differences for continuous and categorical values, respectively.

Model development and validation

For the model development, the K-means clustering algorithm was used. K-means is one of the simplest and most commonly used classes of clustering algorithms. In critical care research, unsupervised machine learning techniques have already been used in several studies, attempting to find homogeneous subgroups within a broad heterogeneous population [26]. This specific algorithm identifies a K number of clusters in a dataset by finding K centroids within the n-dimensional space of clinical features [26].

For feature selection, different sets of candidate variables were tested to assess their ability to produce significantly different mortality probabilities in each cluster using the minimum amount of readily available clinical data. For each set of candidate variables, the optimal number of clusters was determined by comparing models with between 2 and 5 clusters, using the Elbow method [27] and the Calinski-Harabasz index [28]. Information about the methods for selecting number of clusters are provided in the supplemental material.

The following steps were performed for the final model selection: 1) all predictors were assessed for correlation (**eTable 3**); and 2) ten different combinations of the proposed variables were investigated. These combinations were developed based on the perceived clinical importance of each variable and its combinations. All 10 models were tested for the optimal number of clusters and based on both the Elbow method and the Calinski-Harabasz index, as described above. The models were then compared, aiming for the minimum set of variables with high 60-day mortality separation. The description of each model is show in **eTable 4**.

Biological and clinical characteristics of the clusters were evaluated using clinical, laboratory, and (when available) biomarker data to establish subphenotypes [4]. All iterations in model development were done on the training set and the generalizability of the final model was assessed using the validation dataset. K-means clustering analysis is structured to ignore cases with missing data. No assumption was made for missingness and we therefore conducted a complete case analysis. Model development and evaluation was performed using Python version 3.8 and scikit-learn 0.23.1.

Data availability

Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI and SAILS) is publicly available from the NHLBI ARDS Network and data from the ART trial can be requested from study authors.

RESULTS

Participants

Data from 4777 clinical trial patients were considered for inclusion. In total, 4 patients were excluded for having clinical measurements outside plausible range. The remaining 1998 patients from EDEN and FACTT trials were included in the training set, while the 2775 patients from ARMA, ALVEOLI, SAILS, and ART were included in the validation cohort.

Baseline characteristics of the patients in the training and validation sets are presented in **Table 1**. Pneumonia was the prevailing etiology followed by sepsis and aspiration in all trials. Between 29.3% to 72.7% of the patients were receiving vasopressors at the time of randomization. At randomization, PaO_2 / FiO_2 ratio ranged from 112 (75 - 158) to 134 (96 - 185) mmHg, and PEEP from 8 (5 - 10) to 12 (10 - 14) cmH_2O across trials. Mortality at 60 days for the ARDSnet trials ranged from 22.7% to 30.1%, while in the ART trial mortality at 28 days was 58.8%.

Table 1 - Baseline Characteristics and Clinical Outcomes in the Included Trials

	Training set (*n* = 1998)		Validation set (*n* = 2775)			
	EDEN (*n* = 1000)	**FACTT (*n* = 998)**	**ALVEOLI (*n* = 549)**	**ARMA (*n* = 472)**	**ART (*n* = 1010)**	**SAILS (*n* = 744)**
Age, year	52.0 (42.0 - 63.0)	49.0 (38.0 - 60.8)	50.0 (39.0 - 65.0)	50.0 (37.8 - 65.0)	52.0 (36.0 - 64.0)	55.0 (42.0 - 66.0)
Male gender - no. (%)	510 (51.0)	533 (53.4)	302 (55.0)	285 (60.4)	631 (62.5)	365 (49.0)
Etiology - no. (%)						
Pneumonia	650 (65.0)	471 (47.2)	221 (40.3)	145 (30.7)	555 (55.0)	526 (70.7)
Sepsis	147 (14.7)	231 (23.1)	120 (21.9)	125 (26.5)	196 (19.4)	147 (19.8)
Aspiration	96 (9.6)	149 (14.9)	84 (15.3)	72 (15.3)	58 (5.7)	49 (6.6)
Trauma	36 (3.6)	74 (7.4)	45 (8.2)	59 (12.5)	31 (3.1)	6 (0.8)
Other	71 (7.1)	73 (7.3)	79 (14.4)	71 (15.0)	170 (16.8)	16 (2.2)
Severity of Illness*	73.0 (59.0 - 89.0)	78.0 (62.0 - 94.0)	78.0 (64.0 - 93.0)	83.0 (70.0 - 97.0)	63.0 (50.2 – 75.0)	76.0 (61.0 - 92.0)
Vasopressors - no. (%)	489 (48.9)	397 (40.5)	156 (29.3)	147 (31.3)	734 (72.7)	395 (54.2)
Laboratory tests						

White blood cell count, 10^9/L	12.0 (7.8 - 16.7)	11.8 (7.2 - 17.1)	11.6 (7.7 - 15.7)	11.5 (7.5 - 16.2)	---	13.9 (8.7 - 20.0)
Platelets, 10^9/L	169 (108 - 241)	183 (106 - 258)	157 (83 - 247)	135 (80 - 211)	175 (106 - 263)	167 (96 - 247)
Creatinine, mg/dL	1.2 (0.8 - 2.0)	1.0 (0.7 - 1.5)	1.0 (0.7 - 1.7)	1.1 (0.8 - 1.7)	1.3 (0.8 - 2.2)	1.0 (0.7 - 1.7)
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.8 (0.5 - 1.5)	1.0 (0.6 - 2.1)	0.8 (0.4 - 1.5)	0.8 (0.5 - 1.4)
Arterial blood gas						
pH*	7.36 (7.30 - 7.42)	7.37 (7.30 - 7.43)	7.40 (7.34 - 7.44)	7.41 (7.35 - 7.45)	7.28 (7.19 - 7.36)	7.37 (7.31 - 7.42)
PaO_2, mmHg	83 (68 - 108)	79 (67 - 100)	77 (67 - 93)	76.5 (67 - 93)	112 (81 - 155)	83 (69 - 103)
PaO_2 / FiO_2	125 (86 - 178)	118 (80 - 163)	134 (96 - 185)	112 (75 - 158)	112 (81 - 155)	133 (89 - 178)
$PaCO_2$, mmHg	38 (34 - 45)	39 (34 - 45)	38 (33 - 43)	36 (31 - 41)	50 (42 - 62)	39 (34 - 45)
Bicarbonate, mmol/L	21.0 (18.0 - 25.0)	21.0 (17.4 - 25.0)	22.0 (18.0 - 26.0)	22.0 (18.0 - 25.0)	22.9 (19.4 - 26.3)	22.0 (18.0 - 25.0)
Ventilatory variables						
Tidal volume, mL	410 (360 - 470)	450 (400 - 510)	500 (420 - 600)	700 (600 - 750)	350 (308 - 400)	400 (350 - 460)
Per PBW, mL/kg PBW	6.3 (6.0 - 7.3)	7.1 (6.1 - 8.1)	7.9 (6.6 - 9.4)	10.2 (9.0 - 11.3)	5.9 (5.1 - 6.1)	6.2 (6.0 - 7.1)
Plateau pressure, cmH_2O	24.0 (20.0 - 27.0)	26.0 (22.0 - 30.0)	26.0 (22.0 - 31.0)	29.0 (24.8 - 34.0)	26.0 (22.0 - 29.0)	24.0 (19.0 - 28.0)
PEEP, cmH_2O	10 (5 - 12)	10 (5 - 12)	10 (5 - 12)	8 (5 - 10)	12 (10 - 14)	10 (5 - 11)
FiO_2	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.74)	0.70 (0.60 - 1.00)	0.60 (0.40 - 0.70)
Clinical outcomes						
28-day mort. - no. (%)	---	---	---	---	594 (58.8)	---
60-day mort. - no. (%)	227 (22.7)	268 (26.9)	144 (26.2)	141 (30.1)	---	199 (26.7)
90-day mort. - no. (%)	233 (23.3)	283 (28.6)	148 (27.5)	143 (30.8)	---	204 (27.4)
Ventilator-free days, day 28	20.0 (0.0 - 24.0)	17.0 (0.0 - 23.0)	18.0 (0.0 - 24.0)	13.0 (0.0 - 23.0)	0.0 (0.0 - 13.0)	20.0 (0.0 - 25.0)
Ventilator days in survivors	7.0 (4.0 - 13.0)	8.0 (5.0 - 16.0)	8.0 (4.0 - 14.0)	8.0 (4.0 - 15.0)	13.0 (8.0 - 20.0)	6.0 (4.0 - 11.0)

Data are median (quartile 25^{th} - quartile 75^{th}) or N (%)
Abbreviations: 28-day mort. is 28-day mortality, 60-day mort. is 60 days mortality, and 90-day mort. is 90-day mortality.
* Except for ART, that uses SAPS-3, all studies use APACHE-IV

Predictor variables and model selection

The correlation between the 15 variables selected for clustering is shown in **eTable 3**. The strongest correlation was between PEEP and FiO_2 (r = 0.49). The comparison of the 10 models regarding the optimal number of clusters based on both the Elbow method and

the Calinski-Harabasz index is shown in **eFigure 1**. In all models and methods, two clusters were a better fit than a higher number of clusters.

Across the ten models, absolute mortality difference between cluster 1 and cluster 2 ranged from 3.9% to 13.1% for the FACTT study and between 0.1% to 8.1% for EDEN (**Table 2**). The models with the highest 60-day absolute mortality separation between the clusters for each of the two trials in the training set were then further evaluated. Models 6, 5, and 8 were consistently amongst the models with highest separation (**Table 2**). Model 8 was selected for further investigation, as it the fewest variables (**eTable 4**).

Table 2 - Absolute 60-day Mortality Difference Among Clusters per Trial and Model

FACTT trial (*n* = 998)			EDEN trial (*n* = 1000)		
Model	**Patients scored**	**Mortality difference among clusters**	**Model**	**Patients scored**	**Mortality difference among clusters**
6	93.5%	13.1%	7	77.7%	8.1%
2	57.4%	12.5%	8	77.7%	8.1%
5	65.5%	12.2%	6	84.1%	6.7%
8	70.2%	11.6%	5	71.7%	6.5%
7	70.2%	11.4%	9	84.7%	6.1%
1	57.4%	11.2%	3	77.7%	4.4%
4	70.2%	10.6%	4	77.7%	4.0%
9	93.5%	10.4%	2	57.7%	3.9%
3	70.2%	10.1%	10	87.3%	2.8%
10	98.8%	3.9%	1	57.7%	0.1%

Clinical characteristics of each cluster

Based on model 8, only nine clinical and laboratory variables were needed to identify the two distinct clusters in ARDS patients, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. For each variable in the model, opposing measurements could be observed for each cluster

(**Figure 1** and **eFigure 2**). For the ARDSnet trials, the incidence of cluster 1 patients varied from 57.8% (EDEN) to 73.6% (ARMA), and 41.5% of ART patients were part of cluster 1. Across all trials, patients in cluster 2 had higher severity of illness, rate of vasopressor, heart rate, respiratory rate, creatinine, and bilirubin, as well as lower platelets, pH, BUN, and bicarbonate compared to patients in cluster 1 (**eTable 5, 6** and **7**). In addition, 28-, 60-, and 90-day mortality rate was higher in patients in cluster 2 in all trials (**Table 3**). Likewise, for each trial, ventilator-free days at day 28 was lower in patients in cluster 2 compared to cluster 1, and duration of ventilation in survivors was longer in cluster 1.

Table 3 - Clinical Outcomes According to Clusters in Each Trial

	Cluster 1	Cluster 2	Difference (95% CI)	*p* value
Training set				
FACTT	*n* = 407	*n* = 294		
60-day mortality - no. (%)	94 (23.1)	102 (34.7)	11.6% (4.9% to 18.3%)	0.001
90-day mortality - no. (%)	103 (25.4)	106 (36.3)	10.9% (4.1% to 17.8%)	0.002
Ventilator-free days at day 28	19.0 (0.0 - 24.0)	10.0 (0.0 - 21.0)	-9.0 (-11.9 to -6.1)	< 0.001
Duration of ventilation in survivors, days	8.0 (4.0 - 13.0)	10.0 (7.0 - 19.0)	2.0 (0.5 to 3.5)	< 0.001
EDEN	*n* = 449	*n* = 328		
60-day mortality - no. (%)	87 (19.4)	90 (27.4)	8.1% (2.1% to 14.0%)	0.010
90-day mortality - no. (%)	90 (20.0)	93 (28.4)	8.3% (2.3% to 14.3%)	0.009
Ventilator-free days at day 28	21.0 (0.0 - 25.0)	15.0 (0.0 - 22.2)	-6.0 (-8.1 to -3.9)	< 0.001
Duration of ventilation in survivors, days	6.0 (4.0 - 11.0)	8.0 (6.0 - 18.0)	2.0 (0.9 to 3.1)	< 0.001
Validation set				
ALVEOLI	*n* = 336	*n* = 157		
60-day mortality - no. (%)	59 (17.6)	68 (43.3)	25.8% (17.7% to 33.8%)	< 0.001
90-day mortality - no. (%)	60 (18.1)	70 (45.5)	27.3% (19.2% to 35.5%)	< 0.001
Ventilator-free days at day 28	21.0 (4.8 - 25.0)	2.0 (0.0 - 19.0)	-19.0 (-20.8 to -17.2)	< 0.001
Duration of ventilation in survivors, days	7.0 [4.0,13.0]	11.0 (6.0 - 22.2)	4.0 (2.1 to 5.9)	< 0.001
ARMA	*n* = 279	*n* = 100		
60-day mortality - no. (%)	69 (24.8)	42 (42.0)	17.2% (6.9% to 27.5%)	0.002
90-day mortality - no. (%)	70 (25.5)	42 (42.0)	16.5% (6.0% to 26.9%)	0.003
Ventilator-free days at day 28	17.0 (0.0 - 24.0)	2.0 (0.0 - 19.0)	-15.0 (-18.6 to -11.4)	< 0.001
Duration of ventilation in survivors, days	7.0 (4.0 - 13.8)	11.0 (5.0 -18.0)	4.0 (1.5 to 6.5)	0.018
SAILS	*n* = 319	*n* = 188		
60-day mortality - no. (%)	80 (25.1)	60 (31.9)	6.8% (-1.2% to 14.9%)	0.119
90-day mortality - no. (%)	81 (25.4)	63 (33.5)	8.1% (0.0% to 16.3%)	0.063

Ventilator-free days at day 28	21.0 (0.0 - 25.0)	16.0 (0.0 - 23.0)	-5.0 (-7.3 to -2.7)	< 0.001
Duration of ventilation in survivors, days	6.0 (3.0 - 10.0)	8.0 (5.0 - 14.0)	2.0 (0.7 to 3.3)	< 0.001
ART	*n* = 211	*n* = 298		
28-day mortality - no. (%)	81 (38.4)	180 (60.4)	22.0% (13.4% to 30.7%)	< 0.001
Ventilator-free days at day 28	0.0 (0.0 - 17.0)	0.0 (0.0 - 7.8)	-0.0 (-1.0 to 1.0)	< 0.001
Duration of ventilation in survivors, days	12.0 (8.0 - 20.0)	13.5 (8.0 - 20.0)	2.0 (-0.3 to 4.2)	0.570

Data are median (quartile 25th - quartile 75th) or N (%). Difference is mean difference with (95% CI) for binomial variables and median difference with (95% CI) for continuous variables

Abbreviations: CI is confidence interval.

Identification of Subphenotypes

After comparing the clinical characteristics of the clusters, each cluster was assigned to represent a distinct subphenotype of ARDS, with patients in cluster 1 assigned to subphenotype A, and patients in cluster 2 assigned to subphenotype B. Using blood biomarker information available for a subset of patients from both ARMA and ALVEOLI, subphenotype B showed increased levels of pro-inflammatory markers when compared to subphenotype A (**Figure 2** and **eTables 8 and 9**).

DISCUSSION

This study successfully demonstrated that nine easily obtainable clinical variables can identify two distinct ARDS subphenotypes with different clinical and biologic characteristics as well as outcomes across the test and validation cohorts. There was good generalizability amongst diverse populations from multiple validation datasets with temporal and geographical differences.

It is understandable that researchers feel compelled to use as much information as possible to build robust models. This is supportable for two main reasons: (1) the well-known heterogeneity of complex syndromes such as ARDS and (2) the abundance of highly granular clinical data generated by electronic health records (EHRs). It is anticipated that analyzing this vast amount of data will provide new knowledge regarding disease mechanisms by enabling researchers to find plausible hidden patterns within the data [29]. However, this data-heavy approach has the potential drawback of using predictors which are not generally obtained in a time window prior to intervention, or worse yet, using variables that are not part of the routine standard of care for patients. The rationale of using fewer and easy to collect clinical variables is not new in the field of critical care. Prognostic models have already shown that it is indeed feasible to create meaningful models using fewer predictors [30,31].

Our initial choices to define variables commonly found in the EHR at ARDS diagnosis was inspired by a recent report from the World Health Organization (WHO) which showed an enormous discrepancy of medical devices availability in a survey across 135 countries [29]. Recognizing this inconsistency is essential for widespread implementation of machine learning models regardless of varying availability of resources

across countries and health systems [29]. The aim is to provide clinically relevant information within a defined and short time period that might impact the delivery of effective interventions to the right patient population and to as many patients as possible [29].

Recently, Sinha *et al.* developed supervised-learning gradient boosted classifier models trained using 24 or 14 readily available clinical data elements to reproduce biomarker-derived subphenotypes which were previously identified by Calfee *et al.* [17]. Unlike Sinha *et al.*, who predicted previously identified subphenotypes, our study has identified two subphenotypes *de novo* using a small set of clinical variables.

Although the subphenotypes that we have identified and those that have been previously published look similar, our work is distinct from previous studies in several ways. We employed different training and validation datasets and also utilized a different and well-established unsupervised learning technique. Moreover, we utilized a process for selecting predictors which is not comparable to previous studies. Acknowledging these differences is crucial. It would not be unexpected to assume that these deviations would be relevant enough to produce different subphenotypes [32]. However, the clinical, laboratory characteristics, and the clinical outcomes of our subphenotypes show that they are remarkably similar to subphenotypes found in previous papers, regardless of methodological differences.

At this point it is not possible to go beyond this comparative analysis, as there is no gold standard definition of ARDS subphenotypes [32]. Nonetheless, our work does provide robust evidence that ARDS does indeed have two subphenotypes that can be systematically identified, despite major differences in population assessed and

methodological approach used compared with previous studies. It also reinforces that we should continue to explore the underlying biological pathways of such subphenotypes to find responders to new or previously tested therapies.

Our study has several strengths. First, it is the largest cohort of patients that has been studied to develop distinct subphenotypes of ARDS patients. Moreover, our validation cohort included patients from the ART trial, allowing us to validate our model in the contemporaneous population of a large international randomized clinical trial in addition to the ARDSnet studies used in other subphenotyping studies. Second, our subphenotyping model was developed exclusively on the training set and then validated across multiple separate datasets. Nevertheless, similar separation in mortality was seen between the two subphenotypes across all trials. Third, we used the K-means algorithm to identify our subphenotypes, and the results obtained with this technique can be easily interpreted by clinicians and implemented in clinical practice. Lastly, this is the first phenotyping study that has used easily available clinical variables to identify ARDS phenotypes *de novo*, which allows for early identification of these patients in the clinical care at the bedside. Using this algorithm with a small number of routinely collected variables could enable our model to be applied in trials that either retrospectively or prospectively assess interventions targeted to each subphenotype. Future work should analyze previous trials to identify possible differential treatment response for the subphenotypes of ARDS patients identified in this study.

This study also has limitations. First, we have developed our models exclusively on patients enrolled in clinical trials. Due to the strict inclusion and exclusion criteria of these clinical trials, the generalizability of these results needs to be evaluated in

unselected ARDS populations. Although there are clear clinical and biomarker differences between the identified subphenotypes, the model's clinical utility needs to be prospectively validated and further investigated. Additionally, our biomarker analysis is limited to those patients in which the data was made publicly available by the study authors. Lastly, K-means clustering does not handle missing data, and no approach was used to impute missing values. However, the extremely low rate of missingness in our study makes this issue less relevant.

CONCLUSIONS

This study confirms the existence of two distinct subphenotypes in ARDS patients using a novel clustering model on routinely collected clinical data. This work may allow for easier identification of ARDS subphenotypes to facilitate implementation of precision clinical trial enrollment and development of targeted therapies in a variety of settings without the added burdens of biomarker evaluation.

DECLARATIONS

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing Interest: AD, MS, FGZ, ABC, ISM, DMP, LNL declare no relevant financial conflicts of interest. RK, EVA, LB, JO, DR and ROD are employees of Endpoint Health, Inc. ASN reported receiving personal fees from Dräger unrelated to the submitted work.

Ethics Approval: All patients provided informed consent in the original trials. This secondary analysis study was exempt from IRB review because it does not meet the definition of human subject as defined in 45 CFR 46.102. WIRB Work Order #1-1228617-1

Availability of data and material: Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI) is publicly available from the NHLBI ARDS Network (NHLBI ARDS Network) and data from the ART trial can be requested from study authors.

Author Contributions: AD, RK, EVA, LB participated in study design and analysis, drafted, and revised the manuscript, and are the guarantor of the document. MS, DR, JO, FGZ, ABC, ISM, DMP, LNL, and ASN participated in interpretation of data analysis, drafted the manuscript, and revised it for critically important intellectual content. ROD participated in the study design, analysis, interpretation of data analysis, and final revision of the manuscript content.

Twitter: @AbhiduggalMD, @msiuba, @f_g_zampieri, @rod_deliberato, @a_serpaneto, @l_bulgarelli, @endpointhealth

REFERENCES

1 ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, *et al.* Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;**307**:2526–33. doi:10.1001/jama.2012.5669

2 Thille AW, Esteban A, Fernández-Segoviano P, *et al.* Comparison of the Berlin Definition for Acute Respiratory Distress Syndrome with Autopsy. *Am J Respir Crit Care Med* 2013;**187**:761–7. doi:10.1164/rccm.201211-1981OC

3 Reilly J, Calfee C, Christie J. Acute Respiratory Distress Syndrome Phenotypes. *Semin Respir Crit Care Med* 2019;**40**:019–30. doi:10.1055/s-0039-1684049

4 Reddy K, Sinha P, O'Kane CM, *et al.* Subphenotypes in critical care: translation into clinical practice. *The Lancet Respiratory Medicine* 2020;**8**:631–43. doi:10.1016/S2213-2600(20)30124-7

5 Shankar-Hari M, Fan E, Ferguson ND. Acute respiratory distress syndrome (ARDS) phenotyping. *Intensive Care Med* 2019;**45**:516–9. doi:10.1007/s00134-018-5480-6

6 Calfee CS, Delucchi K, Parsons PE, *et al.* Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory Medicine* 2014;**2**:611–20. doi:10.1016/S2213-2600(14)70097-9

7 Famous KR, Delucchi K, Ware LB, *et al.* Acute Respiratory Distress Syndrome Subphenotypes Respond Differently to Randomized Fluid Management Strategy. *Am J Respir Crit Care Med* 2017;**195**:331–8. doi:10.1164/rccm.201603-0645OC

8 Calfee CS, Delucchi KL, Sinha P, *et al.* Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *The Lancet Respiratory Medicine* 2018;**6**:691–8. doi:10.1016/S2213-2600(18)30177-2

9 for the NHLBI ARDS Network, Sinha P, Delucchi KL, *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**:1859–69. doi:10.1007/s00134-018-5378-3

10 Bos LD, Schouten LR, van Vught LA, *et al.* Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax* 2017;**72**:876–83. doi:10.1136/thoraxjnl-2016-209719

11 Ferguson ND, Fan E, Camporota L, *et al.* The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med* 2012;**38**:1573–82. doi:10.1007/s00134-012-2682-1

12 Bellani G, Laffey JG, Pham T, *et al.* Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA* 2016;**315**:788. doi:10.1001/jama.2016.0291

13 Gattinoni L, Vassalli F, Romitti F. Benefits and risks of the P/F approach. *Intensive Care Med* 2018;**44**:2245–7. doi:10.1007/s00134-018-5413-4

14 Matthay MA, Arabi YM, Siegel ER, *et al.* Phenotypes and personalized medicine in the acute respiratory distress syndrome. *Intensive Care Med* 2020;**46**:2136–52. doi:10.1007/s00134-020-06296-9

15 Shankar-Hari M, Rubenfeld GD. Population enrichment for critical care trials: phenotypes and differential outcomes. *Current Opinion in Critical Care* 2019;**25**:489–97. doi:10.1097/MCC.0000000000000641

16 Sinha P, Delucchi KL, McAuley DF, *et al.* Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *The Lancet Respiratory Medicine* 2020;**8**:247–57. doi:10.1016/S2213-2600(19)30369-8

17 Sinha P, Churpek MM, Calfee CS. Machine Learning Classifier Models Can Identify Acute Respiratory Distress Syndrome Phenotypes Using Readily Available Clinical Data. *Am J Respir Crit Care Med* 2020;**202**:996–1004. doi:10.1164/rccm.202002-0347OC

18 Kitsios GD, Yang L, Manatakis DV, *et al.* Host-Response Subphenotypes Offer Prognostic Enrichment in Patients With or at Risk for Acute Respiratory Distress Syndrome*: *Critical Care Medicine* 2019;**47**:1724–34. doi:10.1097/CCM.0000000000004018

19 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome. *N Engl J Med* 2000;**342**:1301–8. doi:10.1056/NEJM200005043421801

20 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Higher versus Lower Positive End-Expiratory Pressures in Patients with the Acute Respiratory Distress Syndrome. *N Engl J Med* 2004;**351**:327–36. doi:10.1056/NEJMoa032193

21 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Pulmonary-Artery versus Central Venous Catheter to Guide Treatment of Acute Lung Injury. *N Engl J Med* 2006;**354**:2213–24. doi:10.1056/NEJMoa061895

22 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Rice TW, Wheeler AP, *et al.* Initial Trophic vs Full

Enteral Feeding in Patients With Acute Lung Injury: The EDEN Randomized Trial. *JAMA: The Journal of the American Medical Association* 2012;**307**:795–803. doi:10.1001/jama.2012.137

23 The National Heart, Lung, and Blood Institute ARDS Clinical Trials Network. Rosuvastatin for Sepsis-Associated Acute Respiratory Distress Syndrome. *N Engl J Med* 2014;**370**:2191–200. doi:10.1056/NEJMoa1401520

24 Writing Group for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) Investigators, Cavalcanti AB, Suzumura ÉA, *et al.* Effect of Lung Recruitment and Titrated Positive End-Expiratory Pressure (PEEP) vs Low PEEP on Mortality in Patients With Acute Respiratory Distress Syndrome: A Randomized Clinical Trial. *JAMA* 2017;**318**:1335. doi:10.1001/jama.2017.14171

25 Yehya N, Harhay MO, Curley MAQ, *et al.* Reappraisal of Ventilator-Free Days in Critical Care Research. *Am J Respir Crit Care Med* 2019;**200**:828–36. doi:10.1164/rccm.201810-2050CP

26 Castela Forte J, Perner A, van der Horst ICC. The use of clustering algorithms in critical care research to unravel patient heterogeneity. *Intensive Care Med* 2019;**45**:1025–8. doi:10.1007/s00134-019-05631-z

27 Ketchen DJ, Shook CL. The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal* 1996;**17**:441–58. doi:https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G

28 Caliński T, Harabasz J. A dendrite method for cluster analysis. *Communications in Statistics* 1974;**3**:1–27. doi:10.1080/03610927408827101

29 Bulgarelli L, Deliberato RO, Johnson AEW. Prediction on critically ill patients: The role of “big data.” *Journal of Critical Care* 2020;**60**:64–8. doi:10.1016/j.jcrc.2020.07.017

30 Johnson AEW, Kramer AA, Clifford GD. A New Severity of Illness Scale Using a Subset of Acute Physiology and Chronic Health Evaluation Data Elements Shows Comparable Predictive Accuracy*: *Critical Care Medicine* 2013;**41**:1711–8. doi:10.1097/CCM.0b013e31828a24fe

31 Deliberato RO, Escudero GG, Bulgarelli L, *et al.* SEVERITAS: An externally validated mortality prediction for critically ill patients in low and middle-income countries. *International Journal of Medical Informatics* 2019;**131**:103959. doi:10.1016/j.ijmedinf.2019.103959

32 DeMerle KM, Angus DC, Baillie JK, *et al.* Sepsis Subclasses: A Framework for Development and Interpretation. *Crit Care Med* Published Online First: 15 February 2021. doi:10.1097/CCM.0000000000004842

FIGURES LEGENDS

Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

Square symbols represent the study with the highest mean z score for each phenotype; Circles represent the study with the lowest mean z score for each phenotype. The colored bands are exclusively to help visualize the opposite trends of the variables on the different clusters; Art.pH: arterial pH; Bicarb: bicarbonate; MAP: mean arterial pressure; Creat: creatinine; Resp.Rate: respiratory rate

Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

For better visualization and due to difference in scales, the values were log-normalized and z-scored. Subphenotypes A and B are shown separately to highlight their differences.

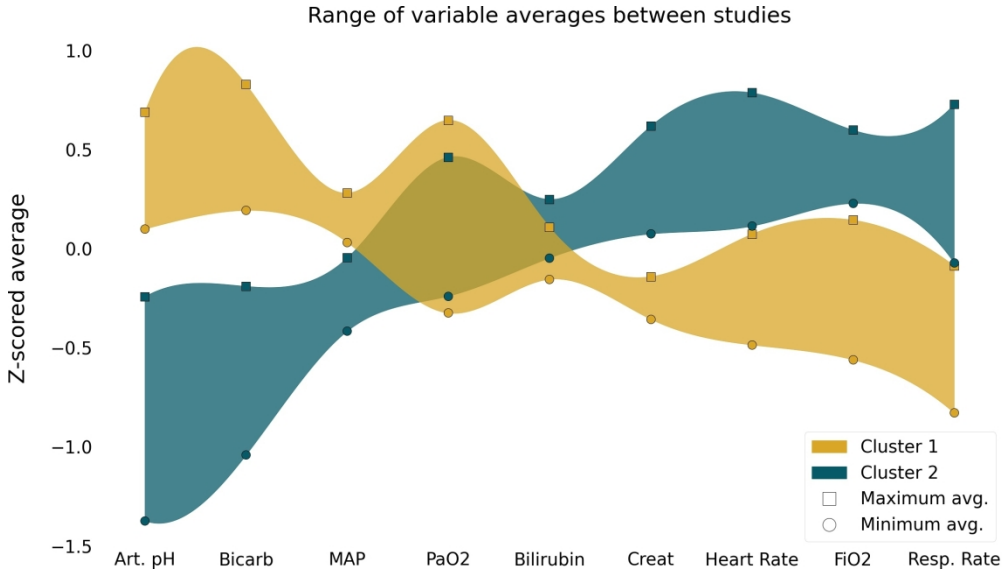


Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

812x457mm (72 x 72 DPI)

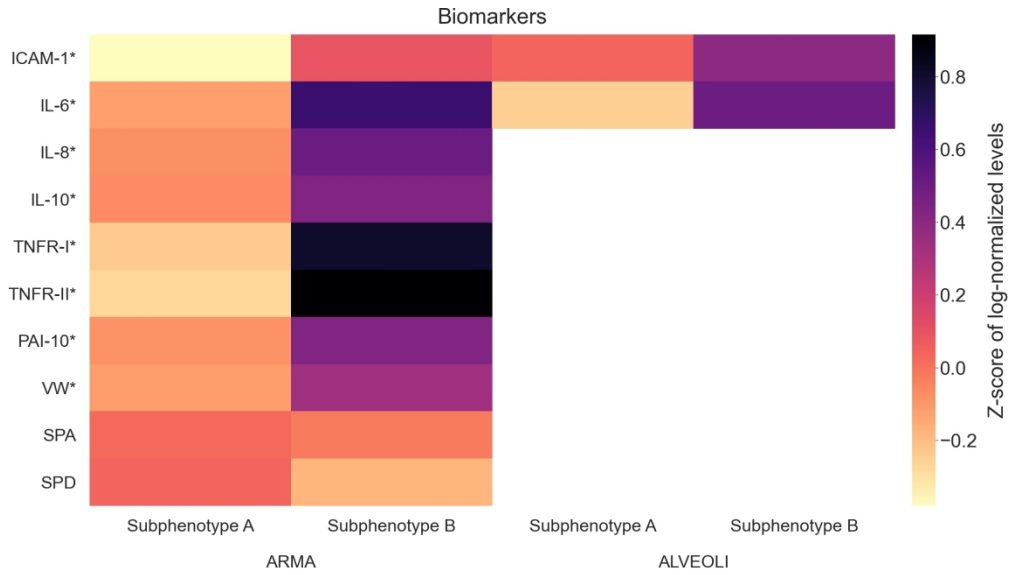


Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

812x457mm (72 x 72 DPI)

Identification of Acute Respiratory Distress Syndrome subphenotypes denovo using routine clinical data: a retrospective analysis of ARDS clinical trials

ONLINE SUPPLEMENT

Additional Methods

Number of clusters

The optimal number of clusters was chosen according to two criteria: (1) Elbow method, by selecting a number of clusters that if further increased will result in only a small increase in performance and possibly cause overfit, hence this number is commonly referenced as to being in the “elbow” of the curve (**eFigure 1**); and (2) Calinski-Harabasz index, consisting of the ratio of *within* to *between* cluster dispersion; higher scores are indication of dense and well separated clusters (**e-Figure 1**).

Ventilator-free days

Ventilator free days for ALVEOLI, EDEN, FACTT, and SAILS were calculated according to the methods outlined by Yehya et al (1). Briefly, patients who died at any time in the 28 days were assigned 0 ventilator-free days. For survivors, the number of ventilator-free days was calculated based on the date of the final successful extubation; reintubations before the final extubation were not counted toward ventilator-free days. All days after a patient was discharged home up to the 28th day with unassisted breathing were assumed to be ventilator-free days.

eTable 1 - Percentage of missing data in the routinely collected variables, closest randomization, on EDEN and FACTT trials.

	EDEN (*n* = 1000)	FACTT (*n* = 999)
Age	0.0	0.0
Gender	0.0	0.0
Arterial pH	2.8	3.9
Bicarbonate	0.2	1.5
Bilirubin	8.1	26.8
Creatinine	0.0	0.0
FiO_2	0.8	0.6
Heart Rate	0.0	0.1
Height	0.1	0.9
Mean Arterial Pressure	12.1	0.8
$PaCO_2$	2.8	3.9
PaO_2	0.2	4.0
Positive end-expiratory pressure	1.0	0.3
Platelets	8.1	6.0
Plateau pressure	32.3	30.9
Respiratory rate	0.6	0.4
Tidal volume	15.3	12.1
Tidal volume per PBW	15.4	12.8

eTable 2 - Plausible physiological ranges for clinical measurements, closest to time of randomization

Variables	Lower Limit	Upper Limit
Age (years)	16	89
Arterial pH	6.65	7.80
Bicarbonate (mEq/L)	1	50
Bilirubin (mg/dL)	0.1	50
Creatinine (mg/dL)	0.1	20
FiO2	0.21	1
Heart Rate (beats per minute)	20	300
Height (cm)	120	220
Mean arterial pressure (mmHg)	10	400
PaCO2 (mmHg)	20	120
PaO2 / FiO2	0	500
PaO2 (mmHg)	30	500
PEEP (cm H20)	0	60
Platelets (thousands)	1	1000
Plateau Pressure (cm H20)	10	50
Respiratory Rate (resp per minute)	1	100
Tidal Volume (cm H20)	100	1400

eTable 3 - Correlation among fifteen routinely collected variables, close to the time of randomization.

	Age	pH	HCO_3	Bili	Creat	FiO_2	Gender	HR	MAP	$PaCO_2$	PaO_2	PEEP	Plat	RR	V_T/PBW
Age	1.00	0.06	-0.04	-0.02	0.11	-0.13	0.00	-0.27	-0.12	-0.11	-0.06	-0.22	0.00	-0.11	0.03
pH	0.06	1.00	0.40	-0.04	-0.16	-0.26	-0.01	-0.18	0.15	-0.39	0.00	-0.20	0.05	-0.21	0.07
HCO_3	-0.04	0.40	1.00	-0.08	-0.28	-0.05	-0.02	-0.18	0.08	0.44	0.02	-0.05	0.15	-0.24	-0.07
Bili	-0.02	-0.04	-0.08	1.00	0.06	-0.03	-0.04	0.01	-0.04	-0.01	0.03	0.01	-0.20	0.04	-0.01
Creat	0.11	-0.16	-0.28	0.06	1.00	-0.04	-0.08	-0.04	-0.01	-0.14	0.00	-0.06	-0.12	0.02	0.00
FiO_2	-0.13	-0.26	-0.05	-0.03	-0.04	1.00	0.03	0.13	-0.06	0.18	0.11	0.49	0.06	0.21	-0.02
Gender	0.00	-0.01	-0.02	-0.04	-0.08	0.03	1.00	-0.03	-0.05	-0.04	-0.06	0.02	0.09	0.09	0.19
HR	-0.27	-0.18	-0.18	0.01	-0.04	0.13	-0.03	1.00	-0.02	0.03	-0.04	0.12	-0.05	0.22	0.08
MAP	-0.12	0.15	0.08	-0.04	-0.01	-0.06	-0.05	-0.02	1.00	-0.03	0.01	-0.01	0.06	-0.04	0.00
$PaCO_2$	-0.11	-0.39	0.44	-0.01	-0.14	0.18	-0.04	0.03	-0.03	1.00	-0.04	0.17	0.11	-0.05	-0.17
PaO_2	-0.06	0.00	0.02	0.03	0.00	0.11	-0.06	-0.04	0.01	-0.04	1.00	-0.09	-0.04	-0.09	0.03
PEEP	-0.22	-0.20	-0.05	0.01	-0.06	0.49	0.02	0.12	-0.01	0.17	-0.09	1.00	0.00	0.33	-0.15
Plat	0.00	0.05	0.15	-0.20	-0.12	0.06	0.09	-0.05	0.06	0.11	-0.04	0.00	1.00	-0.05	0.03
RR	-0.11	-0.21	-0.24	0.04	0.02	0.21	0.09	0.22	-0.04	-0.05	-0.09	0.33	-0.05	1.00	-0.31
V_T/PBW	0.03	0.07	-0.07	-0.01	0.00	-0.02	0.19	0.08	0.00	-0.17	0.03	-0.15	0.03	-0.31	1.00

Data are Pearson correlation coefficients.

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 4 - List of variables in each model assessed

Model	Demographics		Arterial Blood Gases			Laboratory Values				Vital Signs			Ventilator Variables		
	Age	Gender	pH	PaO_2	$PaCO_2$	Creat	Bili	HCO_3	Plat	MAP	RR	HR	FiO2	PEEP	V_T/PBW
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2			X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X		X	X	X	X		
4	X	X	X	X		X	X	X		X	X	X	X		
5			X	X	X	X	X	X	X	X	X	X	X		
6	X	X	X	X		X		X		X	X	X	X		
7			X	X	X	X	X	X		X	X	X	X		
8			X	X		X	X	X		X	X	X	X		
9			X	X	X			X		X	X	X			
10	X	X								X	X	X			

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 5 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Trials in the Training Set

	FACTT			EDEN		
	Cluster 1 (*n* = 407)	**Cluster 2 (*n* = 294)**	***p* value**	**Cluster 1 (*n* = 449)**	**Cluster 2 (*n* = 328)**	***p* value**
Age, year*	50.0 (40.0 - 63.0)	47.0 (36.0 - 58.0)	0.002	53.0 (44.0 - 63.0)	51.0 (41.0 - 62.2)	0.183
Male gender - no. (%)	223 (54.8)	151 (51.4)	0.411	233 (51.9)	168 (51.2)	0.910
Body mass index, kg/m^2	27.5 (23.3 - 32.1)	27.4 (23.0 - 32.7)	0.938	29.1 (24.6 - 34.5)	28.5 (23.4 - 35.1)	0.476
Caucasian - no. (%)	269 (66.1)	177 (60.2)	0.129	349 (81.5)	237 (75.7)	0.067
Etiology - no. (%)			< 0.001			0.003
Pneumonia	201 (49.4)	139 (47.3)		296 (65.9)	217 (66.2)	
Sepsis	78 (19.2)	101 (34.4)		50 (11.1)	60 (18.3)	
Aspiration	67 (16.5)	30 (10.2)		45 (10.0)	27 (8.2)	
Trauma	24 (5.9)	8 (2.7)		24 (5.3)	5 (1.5)	
Other	37 (9.1)	16 (5.4)		34 (7.6)	19 (5.8)	
Prognostic scores						
APACHE III	69.0 (56.0 - 84.0)	91 (76.0 - 105.0)	< 0.001	66.0 (54.0 - 79.0)	84.0 (71.0 - 100.2)	< 0.001
Use of vasopressor - no. (%)	118 (29.5)	189 (64.9)	< 0.001	187 (41.6)	209 (63.7)	< 0.001
Vital signs						
Temperature, ºC	37.5 (36.8 - 38.2)	37.6 (37.0 - 38.4)	0.371	37.3 (36.8 - 37.8)	37.3 (36.7 - 38.1)	0.212
Heart rate, bpm	95.0 (81.0 - 110.0)	114 (102 - 126)	< 0.001	89 (77 - 102)	101 (89 - 116)	< 0.001
Mean arterial Pressure, mmHg	76.0 (68.0 - 88.0)	71.0 (65.0 - 80.8)	< 0.001	77.0 (68.0 - 84.0)	71.0 (66.0 - 80.0)	< 0.001
SpO_2, %	96 (93 - 98)	95 (92 - 97)	< 0.001	96 (94 - 98)	95 (92 - 98)	0.032
Urine output in 24 hours, mL	1785 (1192 - 2853)	1370 (842 - 2446)	< 0.001	1505 (977 - 2250)	1165 (566 - 1816)	< 0.001
Laboratory tests						
Hematocrit, %	30.0 (26.0 - 33.0)	30.0 (24.2 - 35.0)	0.272	30.0 (26.0 - 34.0)	30.0 (26.0 - 35.0)	0.919
White blood cell count, $10^9/L$	11.6 (7.3 - 16.3)	11.7 (5.6 - 17.9)	0.972	11.4 (7.7 - 15.5)	12.7 (7.7 - 19.0)	0.019
Platelets, $10^9/L$	195 (118.5 - 268)	158 (87 - 237)	< 0.001	163 (108 - 241)	164 (103 - 227)	0.552
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.4 (1.0 - 2.0)	< 0.001	1.0 (0.7 - 1.5)	1.6 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.7 (0.5 - 1.3)	0.9 (0.5 - 2.0)	0.003	0.8 (0.5 - 1.3)	0.8 (0.5 - 1.7)	0.128
Arterial blood gas						
pH*	7.41 (7.36 - 7.45)	7.29 (7.23 - 7.35)	< 0.001	7.40 (7.35 - 7.44)	7.30 (7.24 - 7.35)	< 0.001
PaO_2, mmHg	78 (68 - 100)	78 (65 - 99)	0.240	83 (70 - 107)	81 (67 - 107)	0.416
PaO_2 / FiO_2	132 (92 - 173)	89 (65 - 126)	< 0.001	133 (98 - 193)	101 (73 - 162)	< 0.001

$PaCO_2$, mmHg	39 (34 - 44)	38.5 (33 - 47.9)	0.877	38 (34 - 44)	38 (33 - 46)	0.55
Bicarbonate, mmol/L	24.0 (21.0 - 27.0)	17.0 (14.0 - 20.0)	< 0.001	23.0 (21.0 - 26.0)	18.5 (15.0 - 21.0)	< 0.001
Ventilatory variables						
Tidal volume, mL	450 (400 - 530)	450 (382 - 500)	0.009	420 (356 - 487)	400 (350 - 450)	0.032
Per PBW, mL/kg PBW	7.1 (6.3 - 8.4)	7.0 (6.0, 8.0)	0.058	6.3 (6.0 - 7.5)	6.1 (6.0 - 7.3)	0.079
Plateau pressure, cmH_2O	25.0 (20.0 - 29.0)	28.0 (24.0 - 32.0)	< 0.001	23.0 (19.0 - 27.0)	24.0 (21.0 - 28.0)	0.004
PEEP, cmH_2O	8 (5 - 10)	10 (8 - 14)	< 0.001	10 (5 - 10)	10 (8 - 14)	< 0.001
Respiratory rate, breaths/min	22 (18 - 27)	30 (24 - 35)	< 0.001	22 (19 - 26)	30 (25 - 35)	< 0.001
FiO_2	0.50 (0.40 - 0.70)	0.80 (0.60 - 1.00)	< 0.001	0.60 (0.45 - 0.70)	0.80 (0.60 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight.

eTable 6 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	ALVEOLI			ARMA		
	Cluster 1 (*n* = 336)	**Cluster 2 (*n* = 157)**	***p* value**	**Cluster 1 (*n* = 279)**	**Cluster 2 (*n* = 100)**	***p* value**
Age, year*	53.0 (39.0 - 66.2)	46.0 (37.0 - 60.0)	0.007	49.0 (37.0 - 64.0)	47.5 (36.0 - 61.0)	0.180
Male gender - no. (%)	188 (56.0)	86 (54.8)	0.883	169 (60.6)	61 (61.0)	0.965
Body mass index, kg/m^2	27.0 (22.9 - 31.1)	25.2 (21.7 - 30.2)	0.050	25.8 (23.0 - 30.2)	24.4 (21.5 - 29.7)	0.057
Caucasian - no. (%)	263 (78.3)	102 (65.0)	0.002	220 (78.9)	65 (65.0)	0.009
Etiology - no. (%)			0.001			< 0.001
Pneumonia	130 (38.7)	66 (42.0)		83 (29.7)	30 (30.0)	
Sepsis	63 (18.8)	50 (31.8)		64 (22.9)	43 (43.0)	
Aspiration	55 (16.4)	19 (12.1)		44 (15.8)	14 (14.0)	
Trauma	33 (9.8)	5 (3.2)		43 (15.4)	4 (4.0)	
Other	55 (16.4)	17 (10.8)		45 (16.1)	9 (9.0)	
Prognostic scores						
APACHE III	71. (59.0 - 83.0)	93.0 (80.0 - 110.0)	< 0.001	77.0 (66.0 - 90.5)	97.0 (81.8 (110.0)	< 0.001
Use of vasopressor - no. (%)	65 (20.1)	80 (51.3)	< 0.001	77 (27.6)	52 (52.5)	< 0.001
Vital signs						
Temperature, ºC	37.6 (37.1 - 38.2)	37.7 (36.9 - 38.3)	0.778	37.6 (37.1 - 38.1)	37.6 (36.8 - 38.4)	0.803
Heart rate, bpm	97.5 (83.0 - 109)	111.0 (97.0 - 126)	< 0.001	101.0 (89.0 - 112.5)	118 (105.0 - 128.0)	< 0.001
Mean arterial Pressure, mmHg	77.3 (77.0 - 87.3)	73.3 (65.0 - 80.3)	< 0.001	78.0 (70.7 - 88.0)	70.5 (64.9 - 80.4)	< 0.001
SpO_2, %	96 (94 - 97)	95 (92 - 97)	0.005	95 (93 - 98)	95.5 (93 - 97)	0.799
Urine output in 24 hours, mL	2065 (1355 - 3255)	1433 (569 - 2189)	< 0.001	2100 (1375 - 3096)	1525 (816 - 2650)	0.001
Laboratory tests						
Hematocrit, %	31.0 (28.0 - 34.0)	31.0 (27.0 - 35.0)	0.617	30.0 (28.0 - 33.0)	31.0 (28.0 - 34.0)	0.299
White blood cell count, 10^9/L	11.7 (8.1 - 15.3)	10.7 (6.4 - 15.8)	0.166	11.9 (7.7 - 16.7)	9.8 (5.4 - 16.7)	0.057
Platelets, 10^9/L	173 (94 - 266)	141 (57 - 214)	0.001	139 (80 - 212)	125 (72 - 196)	0.260
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.5 (0.9 - 3.0)	< 0.001	1.0 (0.7 - 1.4)	1.8 (1.2 - 3.2)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.9 (0.4 - 1.8)	0.289	1.0 (0.6 - 2.1)	1.1 (0.7 - 2.7)	0.106
Arterial blood gas						
pH*	7.42 (7.38 - 7.45)	7.31 (7.24 - 7.36)	< 0.001	7.42 (7.38 - 7.47)	7.33 (7.28 - 7.37)	< 0.001
PaO_2, mmHg	78 (68 - 93)	74 (65 - 92)	0.082	75 (66 - 91)	81 (68 - 96)	0.106
PaO_2 / FiO_2	149 (109 - 192)	103 (74 - 136)	< 0.001	118 (83 - 160)	99 (68 - 137)	0.006

$PaCO_2$, mmHg	38 (34 - 43)	36 (31 - 42)	0.046	37 (31 - 41)	34 (28.8 - 39.2)	0.003
Bicarbonate, mmol/L	24 (21 - 27)	17 (13 - 20)	< 0.001	23 (20 - 26)	16 (13 - 19)	< 0.001
Ventilatory variables						
Tidal volume, mL	500 (437 - 600)	480 (400 - 572)	0.002	700 (600 - 750)	700 (550 - 700)	0.198
Per PBW, mL/kg PBW	8.0 (6.9 - 9.5)	7.4 (6.2 - 9.2)	0.006	10.1 (9.2 - 11.1)	10.6 (9.0 - 11.4)	0.383
Plateau pressure, cmH_2O	25.0 (21.0 - 30.0)	29.0 (24.0 - 33.0)	< 0.001	29.0 (24.0 - 34.0)	31.0 (27.0 - 36.0)	0.018
PEEP, cmH_2O	10 (5 - 10)	10 (8 - 14)	< 0.001	8 (5 - 10)	10 (5 - 12)	0.150
Respiratory rate, breaths/min	20 (15 - 25)	30 (24 - 35)	< 0.001	18 (14 - 21)	24 (18.8 - 28)	< 0.001
FiO_2	0.50 (0.44 - 0.65)	0.75 (0.60 - 1.00)	< 0.001	0.60 (0.50 - 0.70)	0.70 (0.59 - 0.96)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight.

eTable 7 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	SAILS			ART		
	Cluster 1 (*n* = 319)	**Cluster 2 (*n* = 188)**	***p* value**	**Cluster 1 (*n* = 211)**	**Cluster 2 (*n* = 298)**	***p* value**
Age, year*	57.0 (46.0 - 67.0)	53.5 (39.0 - 65.0)	0.035	54.0 (37.0 - 65.0)	50.0 (35.2 - 61.0)	0.075
Male gender - no. (%)	150 (47.0)	100 (53.2)	0.211	136 (64.5)	181 (60.7)	0.448
Body mass index, kg/m^2	28.5 (23.9 - 34.6)	29.8 (23.2 - 35.1)	0.903	28.8 (24.6 - 35.6)	28.4 (25.0 - 31.7)	0.367
Caucasian - no. (%)	250 (78.4)	140 (74.5)	0.369	---	---	---
Etiology - no. (%)			0.709			0.052
Pneumonia	228 (71.5)	127 (67.6)		113 (53.6)	171 (57.4)	
Sepsis	63 (19.7)	39 (20.7)		38 (18.0)	59 (19.8)	
Aspiration	19 (6.0)	15 (8.0)		13 (6.2)	16 (5.4)	
Trauma	3 (0.9)	1 (0.5)		10 (4.7)	2 (0.7)	
Other	6 (1.9)	6 (3.2)		37 (17.5)	50 (16.8)	
Prognostic scores				---	---	---
APACHE III	70.0 (56.0 - 84.0)	92.0 (75.0 - 105.8)	< 0.001			
SAPS III	---	---	---	62 (50 - 71)	66 (53 - 75)	0.010
Use of vasopressor - no. (%)	150 (47.8)	142 (78.5)	< 0.001	130 (61.6)	242 (81.2)	< 0.001
Vital signs						
Temperature, ºC	37.2 (36.7 - 37.8)	37.3 (36.7 - 38.0)	0.346	---	---	---
Heart rate, bpm	91.0 (80.5 - 103.0)	102.0 (88.8 - 117.0)	< 0.001	90.0 (73.0 - 103.0)	112.0 (97.2 - 126.0)	< 0.001
Mean arterial Pressure, mmHg	78.0 (69.5 - 88.0)	70.0 (63.0 - 78.)	< 0.001	80.0 (73.5 - 89.0)	75.0 (70.0 - 83.0)	< 0.001
SpO_2, %	96 (95 - 99)	96 (93 - 99)	0.270	---	---	---
Urine output in 24 hours, mL	1570 (852 - 2383)	920 (350 - 1665)	< 0.001	---	---	---
Laboratory tests						
Hematocrit, %	31 (27 - 35)	31 (28 - 37)	0.142	---	---	---
White blood cell count, 10^9/L	13.6 (8.5 - 18.1)	15.4 (9.8 - 23.3)	0.009	---	---	---
Platelets, 10^9/L	164 (96 - 238)	131 (80 - 223)	0.032	177 (120 - 292)	169 (90 - 256)	0.048
Creatinine, mg/dL	1.0 (0.7 - 1.5)	1.4 (0.9 - 2.6)	< 0.001	1.0 (0.7 - 1.5)	1.7 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.630	0.6 (0.4 - 1.2)	0.9 (0.4 - 1.7)	0.002
Arterial blood gas						
pH*	7.39 (7.35 - 7.44)	7.31 (7.24 - 7.35)	< 0.001	7.4 (7.3 - 7.4)	7.2 (7.2 - 7.3)	< 0.001
PaO_2, mmHg	82 (68 - 101)	86 (72 - 111.2)	0.112	118 (82 - 158)	104 (78 - 152)	0.065

PaO_2 / FiO_2	139 (98 - 195)	107 (74 - 159)	< 0.001	118 (82 - 158)	104 (78 - 152)	0.065
$PaCO_2$, mmHg	38 (34 - 45)	38 (32 - 44)	0.423	46 (41 - 56)	53 (42 - 65)	< 0.001
Bicarbonate, mmol/L	23 (20 - 26)	17 (14 - 21)	< 0.001	25.2 (22.5 - 28.8)	20.6 (17.8 - 23.4)	< 0.001
Ventilatory variables						
Tidal volume, mL	420 (360 - 480)	400 (340 - 450)	0.016	360 (320 - 400)	350 (300 - 397.8)	0.008
Per PBW, mL/kg PBW	6.4 (6.0 - 7.3)	6.1 (5.9 - 7.0)	0.030	6.0 (5.3 - 6.1)	5.9 (5.1 - 6.1)	0.034
Plateau pressure, cmH_2O	22.0 (18.0 - 27.0)	25.0 (20.0 - 29.0)	0.003	24.0 (21.0 - 28.0)	27.0 (23.0 - 30.0)	< 0.001
PEEP, cmH_2O	8 (5 - 10)	10 (8 - 13)	0.001	10 (10 - 14)	12 (10 - 14)	< 0.001
Respiratory rate, breaths/min	23 (19 - 27)	30 (24 - 35)	< 0.001	24 (20 - 28)	30 (24 - 34)	< 0.001
FiO_2	0.50 (0.40 - 0.60)	0.70 (0.50 - 0.90)	< 0.001	0.70 (0.60 - 0.80)	0.80 (0.70 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight...

eTable 8 - Biomarker levels by study and cluster

	ARMA				ALVEOLI			
	Subphenotype A (*n* = 279)	**Subphenotype B (*n* = 100)**	**Median Difference (95% CI)**	***p* value**	**Subphenotype A (*n* = 336)**	**Subphenotype B (*n* = 157)**	**Median Difference (95% CI)**	***p* value**
ICAM-1	654.0 (399.0 - 959.4)	888.0 (550.0 - 1365.3)	234 (60.3 to 407.8)	0.002	847.9 (585.7 - 1227.1)	1070.4 (748.2 - 1588.8)	219.4 (90.4 to 348.4)	< 0.001
IL-6	214.0 (91.8 - 553.5)	966.0 (291.0 - 2200.0)	749.1 (589.9 to 908.2)	< 0.001	182.5 (85.5 - 435.2)	775.0 (148.0 - 2846.5)	592 (515.5 to 668.6)	< 0.001
PAI-1	65.3 (37.8 - 109.5)	101.7 (50.8 - 291.6)	41 (18.3 to 63.7)	0.001	Not assessed	Not assessed	---	---
IL-8	46.0 (2.0 - 91.0)	106.9 (43.8 - 281.4)	60.9 (35.6 to 86.2)	< 0.001	Not assessed	Not assessed	---	---
IL-10	16.0 (0.0 - 40.3)	47.9 (0.0 - 120.7)	31.9 (20.2 to 43.6)	< 0.001	Not assessed	Not assessed	---	---
TNFR-I	2604.0 (1950.0 - 3777.0)	6897.0 (3622.5 - 12281.5)	4293 (3323.6 to 5262.4)	< 0.001	Not assessed	Not assessed	---	---
TNFR-II	6581.0 (4958.0 - 9658.0)	18611.0 (12262.5 - 35652.0)	12030 (9577.5 to 14482.5)	< 0.001	Not assessed	Not assessed	---	---
SPA	29.0 (11.8 - 68.0)	25.0 (10.5 - 40.0)	-4 (-19.9 to 11.9)	0.398	Not assessed	Not assessed	---	---
SPD	76.0 (36.2 - 145.2)	59.0 (30.0 - 125.0)	-18 (-52.6 to 16.6)	0.254	Not assessed	Not assessed	---	---
VW	308.0 (165.5 - 431.0)	384.0 (246.0 - 549.0)	76 (-26.5 to 178.5)	0.045	Not assessed	Not assessed	---	---

Data are median (quartile 25th - quartile 75th).

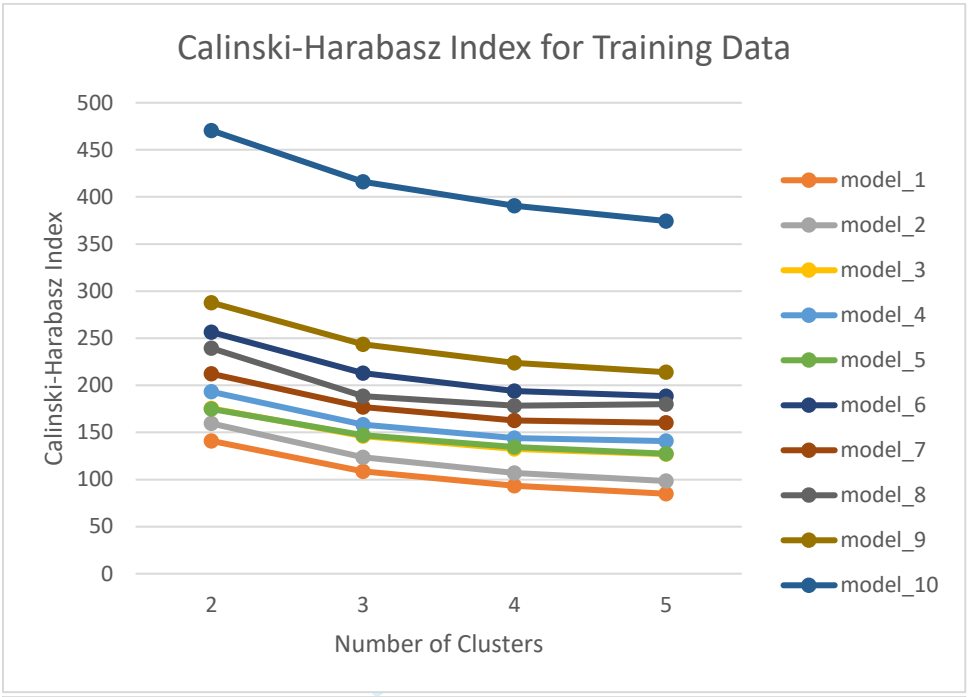
Abbreviations: 95%CI denotes 95% confidence interval, ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

eTable 9 - Percentage of missingness in biomarker levels measured on day of randomization, on ARMA and ALVEOLI trials for patients with an assigned subphenotype

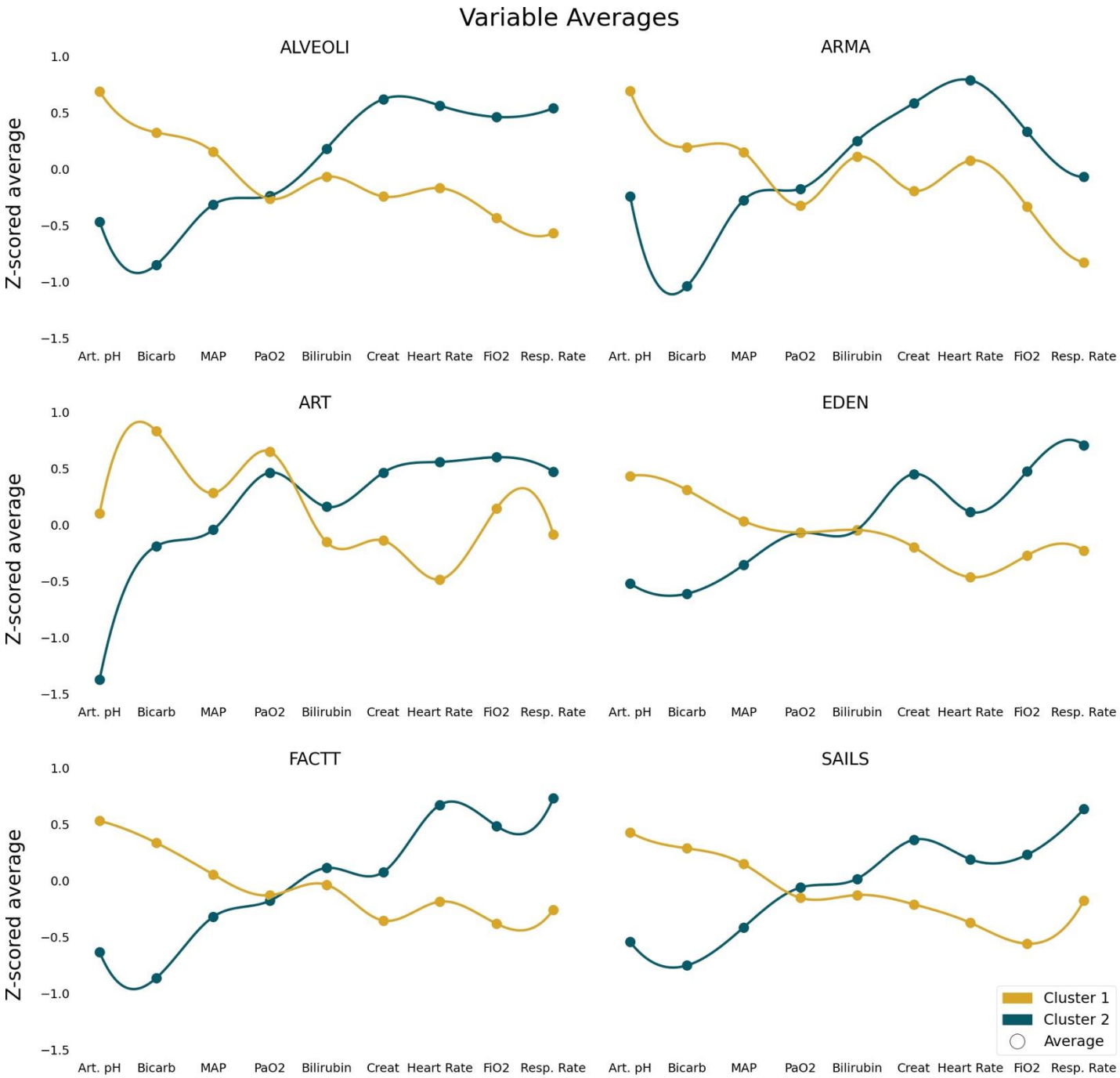
Biomarker	ARMA (*n* = 379)		ALVEOLI (*n* = 493)	
	Subphenotype A	Subphenotype B	Subphenotype A	Subphenotype B
ICAM-1	43%	31%	4%	3%
IL-6	41%	33%	4%	4%
PAI-1	42%	32%	Not assessed	Not assessed
IL-8	41%	33%	Not assessed	Not assessed
IL-10	42%	33%	Not assessed	Not assessed
TNFR-I	68%	61%	Not assessed	Not assessed
TNFR-II	68%	61%	Not assessed	Not assessed
SPA	67%	61%	Not assessed	Not assessed
SPD	67%	61%	Not assessed	Not assessed
VW	67%	61%	Not assessed	Not assessed

Abbreviations: ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

eFigure 1 - Calinski-Harabasz Index and Elbow Method for Each of the 10 Models



eFigure 2 - Variable Averages for Each Study



The circles represent the averages for each variable. The colored lines are exclusively to help visualize the opposite trends of the variables on the different clusters.

Abbreviations: Art. pH is arterial pH, Bicarb is bicarbonate, MAP is mean arterial pressure, Creat is creatinine and Resp. Rate is respiratory rate

Reference:

1. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW. Reappraisal of Ventilator-Free Days in Critical Care Research. Am J Respir Crit Care Med.2019 Oct 1;200(7):828-836. doi: 10.1164/rccm.201810-2050CP.

Reporting checklist for prediction model development/validation.

Based on the TRIPOD guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the TRIPODreporting guidelines, and cite them as:

Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement.

		Reporting Item	Page Number
Title			
	#1	Identify the study as developing and / or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract			
	#2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
	#3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	6
	#3b	Specify the objectives, including whether the study describes the	6

		development or validation of the model or both.	
Methods			
Source of data	#4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8
Source of data	#4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	8
Participants	#5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8
Participants	#5b	Describe eligibility criteria for participants.	8
Participants	#5c	Give details of treatments received, if relevant	8
Outcome	#6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	9
Outcome	#6b	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	#7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured	8
Predictors	#7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	#8	Explain how the study size was arrived at.	8
Missing data	#9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	9
Statistical analysis methods	#10a	If you are developing a prediction model describe how predictors were handled in the analyses.	N/A
Statistical analysis methods	#10b	If you are developing a prediction model, specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	N/A
Statistical analysis methods	#10c	If you are validating a prediction model, describe how the predictions were calculated.	N/A
Statistical analysis	#10d	Specify all measures used to assess model performance and, if relevant,	10

methods		to compare multiple models.	
Statistical analysis methods	#10e	If you are validating a prediction model, describe any model updating (e.g., recalibration) arising from the validation, if done	N/A
Risk groups	#11	Provide details on how risk groups were created, if done.	11
Development vs. validation	#12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	10
Results			
Participants	#13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	12
Participants	#13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	12
Participants	#13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	12
Model development	#14a	If developing a model, specify the number of participants and outcome events in each analysis.	12
Model development	#14b	If developing a model, report the unadjusted association, if calculated between each candidate predictor and outcome.	N/A
Model specification	#15a	If developing a model, present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
Model specification	#15b	If developing a prediction model, explain how to the use it.	N/A
Model performance	#16	Report performance measures (with CIs) for the prediction model.	14
Model-updating	#17	If validating a model, report the results from any model updating, if done (i.e., model specification, model performance).	N/A
Discussion			
Limitations	#18	Discuss any limitations of the study (such as nonrepresentative sample,	19

		few events per predictor, missing data).	
Interpretation	#19a	For validation, discuss the results with reference to performance in the development data, and any other validation data	17
Interpretation	#19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	17
Implications	#20	Discuss the potential clinical use of the model and implications for future research	20
Other information			
Supplementary information	#21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	22
Funding	#22	Give the source of funding and the role of the funders for the present study.	22

This checklist was completed on 07. May 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Identification of Acute Respiratory Distress Syndrome subphenotypes denovo using routine clinical data: a retrospective analysis of ARDS clinical trials

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053297.R1
Article Type:	Original research
Date Submitted by the Author:	26-Oct-2021
Complete List of Authors:	Duggal, Abhijit ; Cleveland Clinic, Department of Critical Care Medicine Kast, Rachel; Endpoint Health Inc, Department of Clinical Data Science Van Ark, Emily; Endpoint Health Inc, Department of Clinical Data Science Bulgarelli, Lucas; Endpoint Health Inc, Department of Clinical Data Science Siuba, Matthew T.; Cleveland Clinic, Department of Critical Care Medicine Osborn, Jeff; Endpoint Health Inc, Department of Clinical Data Science Rey, Diego; Endpoint Health Inc, Department of Clinical Data Science Zampieri, Fernando; HCor Research Institute Cavalcanti, Alexandre ; HCor Research Institute Maia, Israel; Hospital do Coracao Paisani, Denise M; HCor Research Institute Laranjeira, Ligia N; HCor Research Institute Serpa Neto, Ary; Monash University, Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine,; Hospital Israelita Albert Einstein, Critical Care Medicine Deliberato, Rodrigo Octávio; Endpoint Health Inc, Department of Clinical Data Science
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Respiratory medicine
Keywords:	Adult intensive & critical care < INTENSIVE & CRITICAL CARE, Adult intensive & critical care < ANAESTHETICS, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Identification of Acute Respiratory Distress Syndrome subphenotypes de novo using routine clinical data: a retrospective analysis of ARDS clinical trials

Abhijit Duggal MD[1‡], Rachel Kast PhD[2‡], Emily Van Ark PhD[2‡], Lucas Bulgarelli BSc[2‡], Matthew T. Siuba DO[1], Jeff Osborn[2], Diego Rey PhD[2], Fernando G Zampieri MD PhD[3], Alexandre B Cavalcanti MD PhD[3], Israel S Maia MD[3], Denise M Paisani PhD[3], Ligia N Laranjeira[3], Ary Serpa Neto MD MSc PhD,[4,5,6,7,8] Rodrigo Octávio Deliberato MD PhD[2]

‡ Authors contributed equally

1. Department of Critical Care Medicine, Respiratory Institute, Cleveland Clinic, Cleveland, Ohio, USA.
2. Department of Clinical Data Science, Endpoint Health Inc, Palo Alto, California, USA.
3. HCor Research Institute, São Paulo, Brazil
4. Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.
5. Department of Critical Care, Melbourne Medical School, University of Melbourne, Austin Hospital, Melbourne, Australia.
6. Department of Intensive Care, Austin Hospital, Melbourne, Australia.
7. Data Analytics Research and Evaluation (DARE) Centre, Austin Hospital, Melbourne, Australia.
8. Department of Critical Care Medicine, Hospital Israelita Albert Einstein, São Paulo, Brazil.

Correspondence:
Abhijit Duggal MD
Address: 9500 Euclid Ave, L2-330, Cleveland, Ohio, 44195
E-mail: duggala2@ccf.org

Word count (Abstract): 236 words

Word count (Text): 2814 words

Number of figures: 2 figures

Number of tables: 3 tables

Supplementary Material: 01

Key words: Subphenotype, machine learning, ARDS, critical care, clinical data, clustering

ABSTRACT

Objectives: The acute respiratory distress syndrome (ARDS) is a heterogeneous condition, and identification of subphenotypes may help in better risk stratification. Our study objective is to identify ARDS subphenotypes using new simpler methodology and readily available clinical variables.

Setting: This is a retrospective Cohort Study of ARDS trials. Data from the U.S. ARDSNet trials and from the international ART trial.

Participants: 3763 patients from ARDSNet datasets and 1010 patients from the ART dataset.

Primary and secondary outcome measures: The primary outcome was 60-day or 28-day mortality, depending on what was reported in the original trial. K-means cluster analysis was performed to identify subgroups. Sets of candidate variables were tested to assess their ability to produce different probabilities for mortality in each cluster. Clusters were compared to biomarker data, allowing identification of subphenotypes.

Results: Data from 4,773 patients was analyzed. Two subphenotypes (A and B) resulted in optimal separation in the final model, which included nine routinely collected clinical variables, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. Participants in subphenotype B showed increased levels of pro-inflammatory markers, had consistently higher mortality, lower number of ventilator-free days at day 28, and longer duration of ventilation compared to patients in the subphenotype A.

Conclusions: Routinely available clinical data can successfully identify two distinct subphenotypes in adult ARDS patients. This work may facilitate implementation of precision therapy in ARDS clinical trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Largest cohort of patients used to identify subphenotypes of ARDS patients.
- Subphenotypes were validated in the population of a large international ARDS randomized controlled trial.
- Subphenotypes were identified by using only routinely collected clinical data.
- Our use of data exclusively from randomized controlled trials does not prove generalizability to unselected ARDS populations.
- The clinical utility of the subphenotypes has to be validated in a prospective study.

INTRODUCTION

The Berlin definition of acute respiratory distress syndrome (ARDS) encompasses acute hypoxemic respiratory failure due to a wide variety of etiologies [1]. Due to this inclusion of heterogeneous conditions within the syndrome, there are significant clinical and biological differences that make ARDS challenging to treat [2,3]. These differences amongst ARDS patients are associated with variation in risk of disease development and progression [3,4], potentially generating differential responses to treatments and interventions [5–10]. Despite evidence, clinical risk stratification of ARDS patients still solely depends on PaO_2/FiO_2 ratios [11,12], possibly misleading the interpretation of results in clinical trials and clinicians when evaluating treatment options for patients [13].

Therefore, identifying groups of patients who have similar clinical, physiologic, or biomarker traits becomes relevant [6,14] as it can help with stratification of patients producing better targeted therapies and interventions [15]. These different groups can be defined as ARDS subphenotypes [4,14]. Two ARDS subphenotypes have been consistently identified in previous studies [6–10,16–18]. However, these models are complex, and significant barriers exist in their implementation and use in clinical practice. Existing models use up to 40 predictor variables, including biomarkers and other variables that are not readily available at the bedside [6–10,16–18]. These limitations explain the current status quo of ARDS care, where clinicians must depend on the limited prognostic value of PaO_2/FiO_2 ratios instead of biologically distinct subphenotypes.

We hypothesized that the use of a simpler methodology and a small number of easily available clinical variables could identify new ARDS subphenotypes and thus provide the means to allow future implementation of bedside stratification.

METHODS

Data source and participants

We performed a retrospective study using a de-identified dataset pooling data from six randomized clinical trials in patients with ARDS, namely: ARMA, ALVEOLI, FACTT, EDEN, SAILS, and ART [19–24]. Patients in ARMA, ALVEOLI, FACTT, EDEN, and SAILS trials were eligible if they met the American-European consensus for ARDS, including patients with a PaO_2 / FiO_2 ratio < 300 up to 48 hours before enrollment. From 1996 to 2013, these trials enrolled 902, 549, 1000, 1000, and 745 patients, respectively, and tested a variety of interventions [19–23]. Between 2011 and 2017 the international ART study enrolled 1010 adult patients diagnosed with moderate to severe ARDS according to the Berlin definition (PaO_2 / FiO_2 ratio < 200) for less than 72 hours of duration and assessed two different ventilatory strategies [24]. To avoid biases due to high mortality in the high tidal volume group of the ARMA study [19], which has not been standard of care since the beginning of 2000, only 473 patients receiving low tidal volume in that study were included.

Predictors

Six clinical trials were assessed to identify a set of clinical variables recorded closest to time of randomization which were most commonly available across all datasets. The list of potential candidates was then further refined to include only those that are frequently observed in the routine care of ARDS patients at the time of its diagnosis according to judgement provided by ICU physicians who participated in this study. To develop a clustering algorithm for potential rapid translation into clinical use, elements which would not be commonly found in the electronic health records (EHR) at the time of ARDS

diagnosis, such as biomarker levels, ARDS risk factors, organ support apart from mechanical ventilation settings, and severity scores, were excluded from model development. The treatment assignment in the original trials, and clinical outcomes were not considered in the model development.

After all assessment, 16 variables that are routinely collected as part of the usual care and which were uniformly present in all the trials were considered, including: age, gender, arterial pH, PaO_2, $PaCO_2$, bicarbonate, creatinine, bilirubin, platelets, heart rate, respiratory rate, mean arterial pressure, positive end-expiratory pressure (PEEP), plateau pressure, FiO_2, and tidal volume adjusted for predicted body weight (mL/kg PBW). The PBW was calculated as equal to 50 + 0.91 (centimeters of height – 152.4) in males, and 45.5 + 0.91 (centimeters of height – 152.4) in females [18]. These variables were grouped into five domains named demographics, arterial blood gases, laboratory values, vital signs, and ventilatory variables. Plateau pressure was excluded due to a high rate of missingness across the trials included in the training set. Amount of missing data in the training datasets is reported in **eTable 1**.

Outcomes

The primary outcome was 60-day mortality for all ARDSnet trials, and 28-day mortality for ART trial. Secondary outcomes included 90-day mortality, number of ventilator free days at day 28 [25], and the duration of mechanical ventilation in survivors within the first 28 days post enrollment.

Data preparation

Data preprocessing was performed before modeling, and the pooled dataset was assessed for completeness and consistency. Patients with values out of the plausible

physiological range for a specific variable were excluded from the final analysis (described in **eTable 2**). The training dataset was constructed using data from the two largest ARDSnet trials, EDEN and FACTT. The validation dataset was sourced from the four remaining trials: ALVEOLI, ARMA, SAILS, and ART. Means and standard deviations for z-scoring variables were calculated from the training dataset and subsequently applied to the validation data.

Statistical analysis

Baseline and outcome data were presented according to the assigned cluster. Continuous variables were presented as medians with their interquartile ranges and categorical variables as total number and percentage. Proportions were compared using Fisher exact tests and continuous variables were compared using the Wilcoxon rank-sum test. Study outcomes were further compared using the median and mean absolute differences for continuous and categorical values, respectively.

Model development and validation

For the model development, the K-means clustering algorithm was used. K-means is one of the simplest and most used classes of clustering algorithms. In critical care research, unsupervised machine learning techniques have already been used in several studies, attempting to find homogeneous subgroups within a broad heterogeneous population [26]. This specific algorithm identifies a K number of clusters in a dataset by finding K centroids within the n-dimensional space of clinical features [26].

For feature selection, different sets of candidate variables were tested to assess their ability to produce significantly different mortality probabilities in each cluster using the minimum amount of readily available clinical data. For each set of candidate variables,

the optimal number of clusters was determined by comparing models with between 2 and 5 clusters, using the Elbow method [27] and the Calinski-Harabasz index [28]. Information about the methods for selecting the number of clusters are provided in the supplemental material.

The following steps were performed for the final model selection: 1) all predictors were assessed for correlation (**eTable 3**); and 2) ten different combinations of the proposed variables were investigated. These combinations were developed based on the perceived clinical importance of each variable and its combinations. All 10 models were tested for the optimal number of clusters based on both the Elbow method and the Calinski-Harabasz index, as described above. The models were then compared, aiming for the minimum set of variables with high 60-day mortality separation. The description of each model is shown in **eTable 4**.

Biological and clinical characteristics of the clusters were evaluated using clinical, laboratory, and (when available) biomarker data to establish subphenotypes [4]. All iterations in model development were done on the training set and the generalizability of the final model was assessed using the validation dataset. K-means clustering analysis is structured to ignore cases with missing data. No assumption was made for missingness, and we therefore conducted a complete case analysis. Model development and evaluation was performed using Python version 3.8 and scikit-learn 0.23.1.

Patient and public involvement

There was no patient involvement in this study.

Data availability

Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI and SAILS) is publicly available from the NHLBI ARDS Network and data from the ART trial can be requested from study authors.

RESULTS

Participants

Data from 4777 clinical trial patients were considered for inclusion. In total, 4 patients were excluded for having clinical measurements outside plausible range. The remaining 1998 patients from EDEN and FACTT trials were included in the training set, while the 2775 patients from ARMA, ALVEOLI, SAILS, and ART were included in the validation cohort.

Baseline characteristics of the patients in the training and validation sets are presented in **Table 1**. Pneumonia was the prevailing etiology followed by sepsis and aspiration in all trials. Between 29.3% to 72.7% of the patients were receiving vasopressors at the time of randomization. At randomization, PaO_2 / FiO_2 ratio ranged from 112 (75 - 158) to 134 (96 - 185) mmHg, and PEEP from 8 (5 - 10) to 12 (10 - 14) cmH_2O across trials. Mortality at 60 days for the ARDSnet trials ranged from 22.7% to 30.1%, while in the ART trial mortality at 28 days was 58.8%.

Table 1 - Baseline Characteristics and Clinical Outcomes in the Included Trials

	Training set (n = 1998)		Validation set (n = 2775)			
	EDEN (n = 1000)	**FACTT (n = 998)**	**ALVEOLI (n = 549)**	**ARMA (n = 472)**	**ART (n = 1010)**	**SAILS (n = 744)**
Age, year	52.0 (42.0 - 63.0)	49.0 (38.0 - 60.8)	50.0 (39.0 - 65.0)	50.0 (37.8 - 65.0)	52.0 (36.0 - 64.0)	55.0 (42.0 - 66.0)
Male gender - no. (%)	510 (51.0)	533 (53.4)	302 (55.0)	285 (60.4)	631 (62.5)	365 (49.0)
Etiology - no. (%)						
Pneumonia	650 (65.0)	471 (47.2)	221 (40.3)	145 (30.7)	555 (55.0)	526 (70.7)
Sepsis	147 (14.7)	231 (23.1)	120 (21.9)	125 (26.5)	196 (19.4)	147 (19.8)
Aspiration	96 (9.6)	149 (14.9)	84 (15.3)	72 (15.3)	58 (5.7)	49 (6.6)
Trauma	36 (3.6)	74 (7.4)	45 (8.2)	59 (12.5)	31 (3.1)	6 (0.8)
Other	71 (7.1)	73 (7.3)	79 (14.4)	71 (15.0)	170 (16.8)	16 (2.2)
Severity of Illness*	73.0 (59.0 - 89.0)	78.0 (62.0 - 94.0)	78.0 (64.0 - 93.0)	83.0 (70.0 - 97.0)	63.0 (50.2 – 75.0)	76.0 (61.0 - 92.0)
Vasopressors - no. (%)	489 (48.9)	397 (40.5)	156 (29.3)	147 (31.3)	734 (72.7)	395 (54.2)
Laboratory tests						

White blood cell count, 109/L	12.0 (7.8 - 16.7)	11.8 (7.2 - 17.1)	11.6 (7.7 - 15.7)	11.5 (7.5 - 16.2)	---	13.9 (8.7 - 20.0)
Platelets, 109/L	169 (108 - 241)	183 (106 - 258)	157 (83 - 247)	135 (80 - 211)	175 (106 - 263)	167 (96 - 247)
Creatinine, mg/dL	1.2 (0.8 - 2.0)	1.0 (0.7 - 1.5)	1.0 (0.7 - 1.7)	1.1 (0.8 - 1.7)	1.3 (0.8 - 2.2)	1.0 (0.7 - 1.7)
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.8 (0.5 - 1.5)	1.0 (0.6 - 2.1)	0.8 (0.4 - 1.5)	0.8 (0.5 - 1.4)
Arterial blood gas						
pH*	7.36 (7.30 - 7.42)	7.37 (7.30 - 7.43)	7.40 (7.34 - 7.44)	7.41 (7.35 - 7.45)	7.28 (7.19 - 7.36)	7.37 (7.31 - 7.42)
PaO2, mmHg	83 (68 - 108)	79 (67 - 100)	77 (67 - 93)	76.5 (67 - 93)	112 (81 - 155)	83 (69 - 103)
PaO2 / FiO2	125 (86 - 178)	118 (80 - 163)	134 (96 - 185)	112 (75 - 158)	112 (81 - 155)	133 (89 - 178)
PaCO2, mmHg	38 (34 - 45)	39 (34 - 45)	38 (33 - 43)	36 (31 - 41)	50 (42 - 62)	39 (34 - 45)
Bicarbonate, mmol/L	21.0 (18.0 - 25.0)	21.0 (17.4 - 25.0)	22.0 (18.0 - 26.0)	22.0 (18.0 - 25.0)	22.9 (19.4 - 26.3)	22.0 (18.0 - 25.0)
Ventilatory variables						
Tidal volume, mL	410 (360 - 470)	450 (400 - 510)	500 (420 - 600)	700 (600 - 750)	350 (308 - 400)	400 (350 - 460)
Per PBW, mL/kg PBW	6.3 (6.0 - 7.3)	7.1 (6.1 - 8.1)	7.9 (6.6 - 9.4)	10.2 (9.0 - 11.3)	5.9 (5.1 - 6.1)	6.2 (6.0 - 7.1)
Plateau pressure, cmH2O	24.0 (20.0 - 27.0)	26.0 (22.0 - 30.0)	26.0 (22.0 - 31.0)	29.0 (24.8 - 34.0)	26.0 (22.0 - 29.0)	24.0 (19.0 - 28.0)
PEEP, cmH2O	10 (5 - 12)	10 (5 - 12)	10 (5 - 12)	8 (5 - 10)	12 (10 - 14)	10 (5 - 11)
FiO2	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.74)	0.70 (0.60 - 1.00)	0.60 (0.40 - 0.70)
Clinical outcomes						
28-day mort. - no. (%)	---	---	---	---	594 (58.8)	---
60-day mort. - no. (%)	227 (22.7)	268 (26.9)	144 (26.2)	141 (30.1)	---	199 (26.7)
90-day mort. - no. (%)	233 (23.3)	283 (28.6)	148 (27.5)	143 (30.8)	---	204 (27.4)
Ventilator-free days, day 28	20.0 (0.0 - 24.0)	17.0 (0.0 - 23.0)	18.0 (0.0 - 24.0)	13.0 (0.0 - 23.0)	0.0 (0.0 - 13.0)	20.0 (0.0 - 25.0)
Ventilator days in survivors	7.0 (4.0 - 13.0)	8.0 (5.0 - 16.0)	8.0 (4.0 - 14.0)	8.0 (4.0 - 15.0)	13.0 (8.0 - 20.0)	6.0 (4.0 - 11.0)

Data are median (quartile 25th - quartile 75th) or N (%)
Abbreviations: 28-day mort. is 28-day mortality, 60-day mort. is 60 days mortality, and 90-day mort. is 90-day mortality.
* Except for ART, that uses SAPS-3, all studies use APACHE-IV

Predictor variables and model selection

The correlation between the 15 variables selected for clustering is shown in **eTable 3**. The strongest correlation was between PEEP and FiO_2 (r = 0.49). The comparison of the 10 models regarding the optimal number of clusters based on both the Elbow method and the Calinski-Harabasz index is shown in **eFigure 1**. In all models and methods, two clusters were a better fit than a higher number of clusters.

Across the ten models, absolute mortality difference between cluster 1 and cluster 2 ranged from 3.9% to 13.1% for the FACTT study and between 0.1% to 8.1% for EDEN (**eTable 4**). The models with the highest 60-day absolute mortality separation between the clusters for each of the two trials in the training set were then further evaluated. Models 6, 5, and 8 were consistently amongst the models with highest separation (**eTable 4**). Model 8 was selected for further investigation, as it had the fewest variables (**eTable 5**).

Clinical characteristics of each cluster

Based on model 8, only nine clinical and laboratory variables were needed to identify the two distinct clusters in ARDS patients, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. For each variable in the model, opposing measurements could be observed for each cluster (**Figure 1** and **eFigure 2**). For the ARDSnet trials, the incidence of cluster 1 patients varied from 57.8% (EDEN) to 73.6% (ARMA), and 41.5% of ART patients were part of cluster 1. Across all trials, patients in cluster 2 had higher severity of illness, rate of vasopressor, heart rate, respiratory rate, creatinine, and bilirubin, as well as lower platelets, pH, BUN, and bicarbonate compared to patients in cluster 1 (**Table 2, eTables 6** and **7**). In addition, 28-, 60-, and 90-day mortality rate was higher in patients in cluster 2 in all trials (**Table 3**). Likewise, for each trial, the number of ventilator-free days at day 28 was lower in patients in cluster 2 compared to cluster 1, and duration of ventilation in survivors was longer in cluster 1.

Table 2 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Trials in the Training Set						
	FACTT			EDEN		
	Cluster 1 (n = 407)	Cluster 2 (n = 294)	p value	Cluster 1 (n = 449)	Cluster 2 (n = 328)	p value

Age, year*	50.0 (40.0 - 63.0)	47.0 (36.0 - 58.0)	0.002	53.0 (44.0 - 63.0)	51.0 (41.0 - 62.2)	0.183
Male gender - no. (%)	223 (54.8)	151 (51.4)	0.411	233 (51.9)	168 (51.2)	0.910
Body mass index, kg/m2	27.5 (23.3 - 32.1)	27.4 (23.0 - 32.7)	0.938	29.1 (24.6 - 34.5)	28.5 (23.4 - 35.1)	0.476
Caucasian - no. (%)	269 (66.1)	177 (60.2)	0.129	349 (81.5)	237 (75.7)	0.067
Etiology - no. (%)			< 0.001			0.003
Pneumonia	201 (49.4)	139 (47.3)		296 (65.9)	217 (66.2)	
Sepsis	78 (19.2)	101 (34.4)		50 (11.1)	60 (18.3)	
Aspiration	67 (16.5)	30 (10.2)		45 (10.0)	27 (8.2)	
Trauma	24 (5.9)	8 (2.7)		24 (5.3)	5 (1.5)	
Other	37 (9.1)	16 (5.4)		34 (7.6)	19 (5.8)	
Prognostic scores						
APACHE III	69.0 (56.0 - 84.0)	91 (76.0 - 105.0)	< 0.001	66.0 (54.0 - 79.0)	84.0 (71.0 - 100.2)	< 0.001
Use of vasopressor - no. (%)	118 (29.5)	189 (64.9)	< 0.001	187 (41.6)	209 (63.7)	< 0.001
Vital signs						
Temperature, ºC	37.5 (36.8 - 38.2)	37.6 (37.0 - 38.4)	0.371	37.3 (36.8 - 37.8)	37.3 (36.7 - 38.1)	0.212
Heart rate, bpm	95.0 (81.0 - 110.0)	114 (102 - 126)	< 0.001	89 (77 - 102)	101 (89 - 116)	< 0.001
Mean arterial Pressure, mmHg	76.0 (68.0 - 88.0)	71.0 (65.0 - 80.8)	< 0.001	77.0 (68.0 - 84.0)	71.0 (66.0 - 80.0)	< 0.001
SpO2, %	96 (93 - 98)	95 (92 - 97)	< 0.001	96 (94 - 98)	95 (92 - 98)	0.032
Urine output in 24 hours, mL	1785 (1192 - 2853)	1370 (842 - 2446)	< 0.001	1505 (977 - 2250)	1165 (566 - 1816)	< 0.001
Laboratory tests						
Hematocrit, %	30.0 (26.0 - 33.0)	30.0 (24.2 - 35.0)	0.272	30.0 (26.0 - 34.0)	30.0 (26.0 - 35.0)	0.919
White blood cell count, 109/L	11.6 (7.3 - 16.3)	11.7 (5.6 - 17.9)	0.972	11.4 (7.7 - 15.5)	12.7 (7.7 - 19.0)	0.019
Platelets, 109/L	195 (118.5 - 268)	158 (87 - 237)	< 0.001	163 (108 - 241)	164 (103 - 227)	0.552
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.4 (1.0 - 2.0)	< 0.001	1.0 (0.7 - 1.5)	1.6 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.7 (0.5 - 1.3)	0.9 (0.5 - 2.0)	0.003	0.8 (0.5 - 1.3)	0.8 (0.5 - 1.7)	0.128
Arterial blood gas						
pH*	7.41 (7.36 - 7.45)	7.29 (7.23 - 7.35)	< 0.001	7.40 (7.35 - 7.44)	7.30 (7.24 - 7.35)	< 0.001
PaO2, mmHg	78 (68 - 100)	78 (65 - 99)	0.240	83 (70 - 107)	81 (67 - 107)	0.416
PaO2 / FiO2	132 (92 - 173)	89 (65 - 126)	< 0.001	133 (98 - 193)	101 (73 - 162)	< 0.001
PaCO2, mmHg	39 (34 - 44)	38.5 (33 - 47.9)	0.877	38 (34 - 44)	38 (33 - 46)	0.55
Bicarbonate, mmol/L	24.0 (21.0 - 27.0)	17.0 (14.0 - 20.0)	< 0.001	23.0 (21.0 - 26.0)	18.5 (15.0 - 21.0)	< 0.001
Ventilatory variables						
Tidal volume, mL	450 (400 - 530)	450 (382 - 500)	0.009	420 (356 - 487)	400 (350 - 450)	0.032
Per PBW, mL/kg PBW	7.1 (6.3 - 8.4)	7.0 (6.0, 8.0)	0.058	6.3 (6.0 - 7.5)	6.1 (6.0 - 7.3)	0.079
Plateau pressure, cmH2O	25.0 (20.0 - 29.0)	28.0 (24.0 - 32.0)	< 0.001	23.0 (19.0 - 27.0)	24.0 (21.0 - 28.0)	0.004
PEEP, cmH2O	8 (5 - 10)	10 (8 - 14)	< 0.001	10 (5 - 10)	10 (8 - 14)	< 0.001
Respiratory rate, breaths/min	22 (18 - 27)	30 (24 - 35)	< 0.001	22 (19 - 26)	30 (25 - 35)	< 0.001
FiO2	0.50 (0.40 - 0.70)	0.80 (0.60 - 1.00)	< 0.001	0.60 (0.45 - 0.70)	0.80 (0.60 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25th - quartile 75th) or N (%)
Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, VT/PBW denotes tidal volume per predicted body weight.

Table 3 - Clinical Outcomes According to Clusters in Each Trial				
	Cluster 1	**Cluster 2**	**Difference (95% CI)**	**p value**
Training set				
FACTT	n = 407	n = 294		
60-day mortality - no. (%)	94 (23.1)	102 (34.7)	11.6% (4.9% to 18.3%)	0.001
90-day mortality - no. (%)	103 (25.4)	106 (36.3)	10.9% (4.1% to 17.8%)	0.002
Ventilator-free days at day 28	19.0 (0.0 - 24.0)	10.0 (0.0 - 21.0)	-9.0 (-11.9 to -6.1)	< 0.001
Duration of ventilation in survivors, days	8.0 (4.0 - 13.0)	10.0 (7.0 - 19.0)	2.0 (0.5 to 3.5)	< 0.001
EDEN	n = 449	n = 328		
60-day mortality - no. (%)	87 (19.4)	90 (27.4)	8.1% (2.1% to 14.0%)	0.010
90-day mortality - no. (%)	90 (20.0)	93 (28.4)	8.3% (2.3% to 14.3%)	0.009
Ventilator-free days at day 28	21.0 (0.0 - 25.0)	15.0 (0.0 - 22.2)	-6.0 (-8.1 to -3.9)	< 0.001
Duration of ventilation in survivors, days	6.0 (4.0 - 11.0)	8.0 (6.0 - 18.0)	2.0 (0.9 to 3.1)	< 0.001
Validation set				
ALVEOLI	n = 336	n = 157		
60-day mortality - no. (%)	59 (17.6)	68 (43.3)	25.8% (17.7% to 33.8%)	< 0.001
90-day mortality - no. (%)	60 (18.1)	70 (45.5)	27.3% (19.2% to 35.5%)	< 0.001
Ventilator-free days at day 28	21.0 (4.8 - 25.0)	2.0 (0.0 - 19.0)	-19.0 (-20.8 to -17.2)	< 0.001
Duration of ventilation in survivors, days	7.0 [4.0,13.0]	11.0 (6.0 - 22.2)	4.0 (2.1 to 5.9)	< 0.001
ARMA	n = 279	n = 100		
60-day mortality - no. (%)	69 (24.8)	42 (42.0)	17.2% (6.9% to 27.5%)	0.002
90-day mortality - no. (%)	70 (25.5)	42 (42.0)	16.5% (6.0% to 26.9%)	0.003
Ventilator-free days at day 28	17.0 (0.0 - 24.0)	2.0 (0.0 - 19.0)	-15.0 (-18.6 to -11.4)	< 0.001
Duration of ventilation in survivors, days	7.0 (4.0 - 13.8)	11.0 (5.0 -18.0)	4.0 (1.5 to 6.5)	0.018
SAILS	n = 319	n = 188		
60-day mortality - no. (%)	80 (25.1)	60 (31.9)	6.8% (-1.2% to 14.9%)	0.119
90-day mortality - no. (%)	81 (25.4)	63 (33.5)	8.1% (0.0% to 16.3%)	0.063
Ventilator-free days at day 28	21.0 (0.0 - 25.0)	16.0 (0.0 - 23.0)	-5.0 (-7.3 to -2.7)	< 0.001
Duration of ventilation in survivors, days	6.0 (3.0 - 10.0)	8.0 (5.0 - 14.0)	2.0 (0.7 to 3.3)	< 0.001
ART	n = 211	n = 298		
28-day mortality - no. (%)	81 (38.4)	180 (60.4)	22.0% (13.4% to 30.7%)	< 0.001
Ventilator-free days at day 28	0.0 (0.0 - 17.0)	0.0 (0.0 - 7.8)	-0.0 (-1.0 to 1.0)	< 0.001
Duration of ventilation in survivors, days	12.0 (8.0 - 20.0)	13.5 (8.0 - 20.0)	2.0 (-0.3 to 4.2)	0.570

Data are median (quartile 25th - quartile 75th) or N (%). Difference is mean difference with (95% CI) for binomial variables and median difference with (95% CI) for continuous variables
Abbreviations: CI is the confidence interval.

Identification of Subphenotypes

After comparing the clinical characteristics of the clusters, each cluster was assigned to represent a distinct subphenotype of ARDS, with patients in cluster 1 assigned to

subphenotype A, and patients in cluster 2 assigned to subphenotype B. Using blood biomarker information available for a subset of patients from both ARMA and ALVEOLI, subphenotype B showed increased levels of pro-inflammatory markers when compared to subphenotype A (**Figure 2** and **eTables 8 and 9**).

DISCUSSION

This study successfully demonstrated that nine easily obtainable clinical variables: arterial pH, partial O2 pressure, creatinine, bilirubin, bicarbonate, mean arterial pressure, heart rate, respiratory rate, and FiO2 at the time of study enrollment can identify two distinct ARDS subphenotypes with different clinical and biologic characteristics as well as outcomes across the test and validation cohorts. There was good generalizability amongst diverse populations from multiple validation datasets with temporal and geographical differences.

It is understandable that researchers feel compelled to use as much information as possible to build robust models. This is supportable for two main reasons: (1) the well-known heterogeneity of complex syndromes such as ARDS and (2) the abundance of highly granular clinical data generated by electronic health records (EHRs). It is anticipated that analyzing this vast amount of data will provide new knowledge regarding disease mechanisms by enabling researchers to find plausible hidden patterns within the data [29]. However, this data-heavy approach has the potential drawback of using predictors which are not generally obtained in a time window prior to intervention, or worse yet, using variables that are not part of the routine standard of care for patients. The rationale of using fewer and easy to collect clinical variables is not new in the field of critical care. Prognostic models have already shown that it is indeed feasible to create meaningful models using fewer predictors [30,31].

Our initial choices to define variables commonly found in the EHR at ARDS diagnosis was inspired by a recent report from the World Health Organization (WHO) which showed an enormous discrepancy of medical devices availability in a survey across

135 countries [29]. Recognizing this inconsistency is essential for widespread implementation of machine learning models regardless of varying availability of resources across countries and health systems [29]. The aim is to provide clinically relevant information within a defined and short period that might impact the delivery of effective interventions to the right patient population and to as many patients as possible [29].

Recently, Sinha *et al.* developed supervised-learning gradient boosted classifier models trained using 24 or 14 readily available clinical data elements to reproduce biomarker-derived subphenotypes which were previously identified by Calfee *et al.* [17]. Unlike Sinha *et al.*, who predicted previously identified subphenotypes, our study has identified two subphenotypes *de novo* using a small set of clinical variables.

Although the subphenotypes that we have identified and those that have been previously published look similar, our work is distinct from previous studies in several ways. We employed different training and validation datasets as well as a different and well-established unsupervised learning technique. Moreover, we utilized a process for selecting predictors which is not comparable to previous studies. Acknowledging these differences is crucial. It would not be unexpected to assume that these deviations would be relevant enough to produce different subphenotypes [32]. However, the clinical, laboratory characteristics, and the clinical outcomes of our subphenotypes show that they are remarkably similar to subphenotypes found in previous papers, regardless of methodological differences.

At this point it is not possible to go beyond this comparative analysis, as there is no gold standard definition of ARDS subphenotypes [32]. Nonetheless, our work does provide robust evidence that ARDS does indeed have two subphenotypes that can be

systematically identified, despite major differences in population assessed and methodological approach used compared with previous studies. It also reinforces that we should continue to explore the underlying biological pathways of such subphenotypes to find responders to new or previously tested therapies.

Our study has several strengths. First, it is the largest cohort of patients that has been studied to develop distinct subphenotypes of ARDS patients. Moreover, our validation cohort included patients from the ART trial, allowing us to validate our model in the contemporaneous population of a large international randomized clinical trial in addition to the ARDSnet studies used in other subphenotyping studies. Second, our subphenotyping model was developed exclusively on the training set and then validated across multiple separate datasets. Nevertheless, similar separation in mortality was seen between the two subphenotypes across all trials. Third, we used the K-means algorithm to identify our subphenotypes, and the results obtained with this technique can be easily interpreted by clinicians and implemented in clinical practice. Lastly, this is the first phenotyping study that has used easily available clinical variables to identify ARDS phenotypes *de novo*, which allows for early identification of these patients in the clinical care at the bedside. Using this algorithm with a small number of routinely collected variables could enable our model to be applied in trials that either retrospectively or prospectively assess interventions targeted to each subphenotype.

This study also has limitations. First, we have developed our models exclusively on patients enrolled in clinical trials. Due to the strict inclusion and exclusion criteria of these clinical trials, the generalizability of these results needs to be evaluated in unselected ARDS populations. Although there are clear clinical and biomarker differences

between the identified subphenotypes, the model's clinical utility needs to be prospectively validated and further investigated. Additionally, our biomarker analysis is limited to those patients in which the data was made publicly available by the study authors, but future collection of biomarker data in a prospective study will allow more robust understanding of the underlying biology and validation of the subphenotype model. Also, K-means clustering does not handle missing data, and no approach was used to impute missing values. However, the extremely low rate of missingness in our study makes this issue less relevant. Lastly, future work should analyze previous trials to identify possible differential treatment responses for the subphenotypes of ARDS patients identified in this study.

CONCLUSIONS

This study confirms the existence of two distinct subphenotypes in ARDS patients using a novel clustering model on routinely collected clinical data. This work may allow for easier identification of ARDS subphenotypes to facilitate implementation of precision clinical trial enrollment and development of targeted therapies in a variety of settings without the added burdens of biomarker evaluation.

DECLARATIONS

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing Interest: AD, MS, FGZ, ABC, ISM, DMP, LNL declare no relevant financial conflicts of interest. RK, EVA, LB, JO, DR and ROD are employees of Endpoint Health, Inc. ASN reported receiving personal fees from Dräger unrelated to the submitted work.

Ethics Approval: All patients provided informed consent in the original trials. This secondary analysis study was exempt from IRB review because it does not meet the definition of human subject as defined in 45 CFR 46.102. WIRB Work Order #1-1228617-1

Availability of data and material: Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI) is publicly available from the NHLBI ARDS Network (NHLBI ARDS Network) and data from the ART trial can be requested from study authors.

Author Contributions: AD, RK, EVA, LB participated in study design and analysis, drafted, and revised the manuscript, and are the guarantor of the document. MS, DR, JO, FGZ, ABC, ISM, DMP, LNL, and ASN participated in interpretation of data analysis, drafted the manuscript, and revised it for critically important intellectual content. ROD participated in the study design, analysis, interpretation of data analysis, and final revision of the manuscript content.

Twitter: @msiuba, @f_g_zampieri, @rod_deliberato, @a_serpaneto, @l_bulgarelli, @endpointhealth

REFERENCES

1 ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, *et al.* Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;**307**:2526–33. doi:10.1001/jama.2012.5669

2 Thille AW, Esteban A, Fernández-Segoviano P, *et al.* Comparison of the Berlin Definition for Acute Respiratory Distress Syndrome with Autopsy. *Am J Respir Crit Care Med* 2013;**187**:761–7. doi:10.1164/rccm.201211-1981OC

3 Reilly J, Calfee C, Christie J. Acute Respiratory Distress Syndrome Phenotypes. *Semin Respir Crit Care Med* 2019;**40**:019–30. doi:10.1055/s-0039-1684049

4 Reddy K, Sinha P, O'Kane CM, *et al.* Subphenotypes in critical care: translation into clinical practice. *The Lancet Respiratory Medicine* 2020;**8**:631–43. doi:10.1016/S2213-2600(20)30124-7

5 Shankar-Hari M, Fan E, Ferguson ND. Acute respiratory distress syndrome (ARDS) phenotyping. *Intensive Care Med* 2019;**45**:516–9. doi:10.1007/s00134-018-5480-6

6 Calfee CS, Delucchi K, Parsons PE, *et al.* Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory Medicine* 2014;**2**:611–20. doi:10.1016/S2213-2600(14)70097-9

7 Famous KR, Delucchi K, Ware LB, *et al.* Acute Respiratory Distress Syndrome Subphenotypes Respond Differently to Randomized Fluid Management Strategy. *Am J Respir Crit Care Med* 2017;**195**:331–8. doi:10.1164/rccm.201603-0645OC

8 Calfee CS, Delucchi KL, Sinha P, *et al.* Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *The Lancet Respiratory Medicine* 2018;**6**:691–8. doi:10.1016/S2213-2600(18)30177-2

9 for the NHLBI ARDS Network, Sinha P, Delucchi KL, *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**:1859–69. doi:10.1007/s00134-018-5378-3

10 Bos LD, Schouten LR, van Vught LA, *et al.* Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax* 2017;**72**:876–83. doi:10.1136/thoraxjnl-2016-209719

11 Ferguson ND, Fan E, Camporota L, *et al.* The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med* 2012;**38**:1573–82. doi:10.1007/s00134-012-2682-1

12 Bellani G, Laffey JG, Pham T, *et al.* Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA* 2016;**315**:788. doi:10.1001/jama.2016.0291

13 Gattinoni L, Vassalli F, Romitti F. Benefits and risks of the P/F approach. *Intensive Care Med* 2018;**44**:2245–7. doi:10.1007/s00134-018-5413-4

14 Matthay MA, Arabi YM, Siegel ER, *et al.* Phenotypes and personalized medicine in the acute respiratory distress syndrome. *Intensive Care Med* 2020;**46**:2136–52. doi:10.1007/s00134-020-06296-9

15 Shankar-Hari M, Rubenfeld GD. Population enrichment for critical care trials: phenotypes and differential outcomes. *Current Opinion in Critical Care* 2019;**25**:489–97. doi:10.1097/MCC.0000000000000641

16 Sinha P, Delucchi KL, McAuley DF, *et al.* Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *The Lancet Respiratory Medicine* 2020;**8**:247–57. doi:10.1016/S2213-2600(19)30369-8

17 Sinha P, Churpek MM, Calfee CS. Machine Learning Classifier Models Can Identify Acute Respiratory Distress Syndrome Phenotypes Using Readily Available Clinical Data. *Am J Respir Crit Care Med* 2020;**202**:996–1004. doi:10.1164/rccm.202002-0347OC

18 Kitsios GD, Yang L, Manatakis DV, *et al.* Host-Response Subphenotypes Offer Prognostic Enrichment in Patients With or at Risk for Acute Respiratory Distress Syndrome*: *Critical Care Medicine* 2019;**47**:1724–34. doi:10.1097/CCM.0000000000004018

19 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome. *N Engl J Med* 2000;**342**:1301–8. doi:10.1056/NEJM200005043421801

20 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Higher versus Lower Positive End-Expiratory Pressures in Patients with the Acute Respiratory Distress Syndrome. *N Engl J Med* 2004;**351**:327–36. doi:10.1056/NEJMoa032193

21 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Pulmonary-Artery versus Central Venous Catheter to Guide Treatment of Acute Lung Injury. *N Engl J Med* 2006;**354**:2213–24. doi:10.1056/NEJMoa061895

22 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Rice TW, Wheeler AP, *et al.* Initial Trophic vs Full

Enteral Feeding in Patients With Acute Lung Injury: The EDEN Randomized Trial. *JAMA: The Journal of the American Medical Association* 2012;**307**:795–803. doi:10.1001/jama.2012.137

23 The National Heart, Lung, and Blood Institute ARDS Clinical Trials Network. Rosuvastatin for Sepsis-Associated Acute Respiratory Distress Syndrome. *N Engl J Med* 2014;**370**:2191–200. doi:10.1056/NEJMoa1401520

24 Writing Group for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) Investigators, Cavalcanti AB, Suzumura ÉA, *et al.* Effect of Lung Recruitment and Titrated Positive End-Expiratory Pressure (PEEP) vs Low PEEP on Mortality in Patients With Acute Respiratory Distress Syndrome: A Randomized Clinical Trial. *JAMA* 2017;**318**:1335. doi:10.1001/jama.2017.14171

25 Yehya N, Harhay MO, Curley MAQ, *et al.* Reappraisal of Ventilator-Free Days in Critical Care Research. *Am J Respir Crit Care Med* 2019;**200**:828–36. doi:10.1164/rccm.201810-2050CP

26 Castela Forte J, Perner A, van der Horst ICC. The use of clustering algorithms in critical care research to unravel patient heterogeneity. *Intensive Care Med* 2019;**45**:1025–8. doi:10.1007/s00134-019-05631-z

27 Ketchen DJ, Shook CL. The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal* 1996;**17**:441–58. doi:https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G

28 Caliński T, Harabasz J. A dendrite method for cluster analysis. *Communications in Statistics* 1974;**3**:1–27. doi:10.1080/03610927408827101

29 Bulgarelli L, Deliberato RO, Johnson AEW. Prediction on critically ill patients: The role of “big data.” *Journal of Critical Care* 2020;**60**:64–8. doi:10.1016/j.jcrc.2020.07.017

30 Johnson AEW, Kramer AA, Clifford GD. A New Severity of Illness Scale Using a Subset of Acute Physiology and Chronic Health Evaluation Data Elements Shows Comparable Predictive Accuracy*: *Critical Care Medicine* 2013;**41**:1711–8. doi:10.1097/CCM.0b013e31828a24fe

31 Deliberato RO, Escudero GG, Bulgarelli L, *et al.* SEVERITAS: An externally validated mortality prediction for critically ill patients in low and middle-income countries. *International Journal of Medical Informatics* 2019;**131**:103959. doi:10.1016/j.ijmedinf.2019.103959

32 DeMerle KM, Angus DC, Baillie JK, *et al.* Sepsis Subclasses: A Framework for Development and Interpretation. *Crit Care Med* Published Online First: 15 February 2021. doi:10.1097/CCM.0000000000004842

FIGURES LEGENDS

Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

Square symbols represent the study with the highest mean z score for each phenotype; Circles represent the study with the lowest mean z score for each phenotype. The colored bands are exclusively to help visualize the opposite trends of the variables on the different clusters; Art.pH: arterial pH; Bicarb: bicarbonate; MAP: mean arterial pressure; Creat: creatinine; Resp.Rate: respiratory rate

Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

For better visualization and due to difference in scales, the values were log-normalized and z-scored. Subphenotypes A and B are shown separately to highlight their differences.

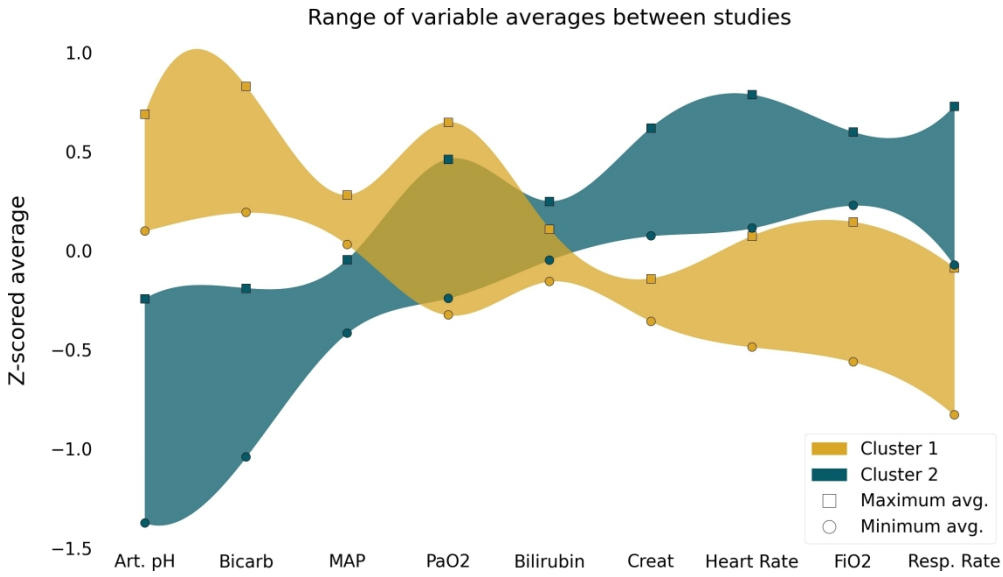


Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

812x457mm (72 x 72 DPI)

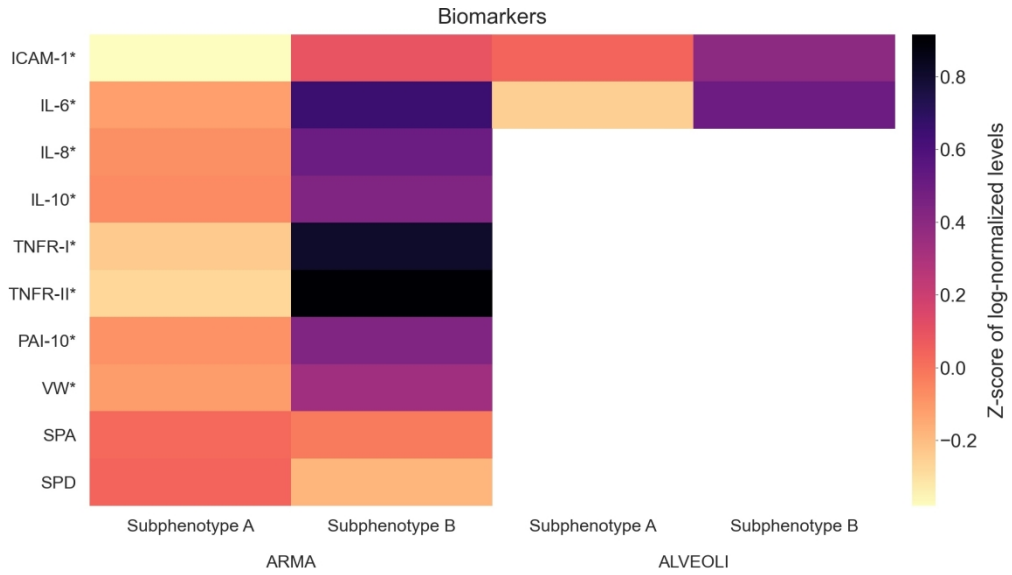


Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

812x457mm (72 x 72 DPI)

Identification of Acute Respiratory Distress Syndrome subphenotypes de novo using routine clinical data: a retrospective analysis of ARDS clinical trials

ONLINE SUPPLEMENT

Additional Methods

Number of clusters

The optimal number of clusters was chosen according to two criteria: (1) Elbow method, by selecting a number of clusters that if further increased will result in only a small increase in performance and possibly cause overfit, hence this number is commonly referenced as to being in the "elbow" of the curve (**eFigure 1**); and (2) Calinski-Harabasz index, consisting of the ratio of *within* to *between* cluster dispersion; higher scores are indication of dense and well separated clusters (**e-Figure 1**).

Ventilator-free days

Ventilator free days for ALVEOLI, EDEN, FACTT, and SAILS were calculated according to the methods outlined by Yehya et al (1). Briefly, patients who died at any time in the 28 days were assigned 0 ventilator-free days. For survivors, the number of ventilator-free days was calculated based on the date of the final successful extubation; reintubations before the final extubation were not counted toward ventilator-free days. All days after a patient was discharged home up to the 28th day with unassisted breathing were assumed to be ventilator-free days.

eTable 1 - Percentage of missing data in the routinely collected variables, closest randomization, on EDEN and FACTT trials.

	EDEN (*n* = 1000)	FACTT (*n* = 999)
Age	0.0	0.0
Gender	0.0	0.0
Arterial pH	2.8	3.9
Bicarbonate	0.2	1.5
Bilirubin	8.1	26.8
Creatinine	0.0	0.0
FiO_2	0.8	0.6
Heart Rate	0.0	0.1
Height	0.1	0.9
Mean Arterial Pressure	12.1	0.8
$PaCO_2$	2.8	3.9
PaO_2	0.2	4.0
Positive end-expiratory pressure	1.0	0.3
Platelets	8.1	6.0
Plateau pressure	32.3	30.9
Respiratory rate	0.6	0.4
Tidal volume	15.3	12.1
Tidal volume per PBW	15.4	12.8

eTable 2 - Plausible physiological ranges for clinical measurements, closest to time of randomization

Variables	Lower Limit	Upper Limit
Age (years)	16	89
Arterial pH	6.65	7.80
Bicarbonate (mEq/L)	1	50
Bilirubin (mg/dL)	0.1	50
Creatinine (mg/dL)	0.1	20
FiO2	0.21	1
Heart Rate (beats per minute)	20	300
Height (cm)	120	220
Mean arterial pressure (mmHg)	10	400
PaCO2 (mmHg)	20	120
PaO2 / FiO2	0	500
PaO2 (mmHg)	30	500
PEEP (cm H20)	0	60
Platelets (thousands)	1	1000
Plateau Pressure (cm H20)	10	50
Respiratory Rate (resp per minute)	1	100
Tidal Volume (cm H20)	100	1400

eTable 3 - Correlation among fifteen routinely collected variables, close to the time of randomization.

	Age	pH	HCO_3	Bili	Creat	FiO_2	Gender	HR	MAP	$PaCO_2$	PaO_2	PEEP	Plat	RR	V_T/PBW
Age	1.00	0.06	-0.04	-0.02	0.11	-0.13	0.00	-0.27	-0.12	-0.11	-0.06	-0.22	0.00	-0.11	0.03
pH	0.06	1.00	0.40	-0.04	-0.16	-0.26	-0.01	-0.18	0.15	-0.39	0.00	-0.20	0.05	-0.21	0.07
HCO_3	-0.04	0.40	1.00	-0.08	-0.28	-0.05	-0.02	-0.18	0.08	0.44	0.02	-0.05	0.15	-0.24	-0.07
Bili	-0.02	-0.04	-0.08	1.00	0.06	-0.03	-0.04	0.01	-0.04	-0.01	0.03	0.01	-0.20	0.04	-0.01
Creat	0.11	-0.16	-0.28	0.06	1.00	-0.04	-0.08	-0.04	-0.01	-0.14	0.00	-0.06	-0.12	0.02	0.00
FiO_2	-0.13	-0.26	-0.05	-0.03	-0.04	1.00	0.03	0.13	-0.06	0.18	0.11	0.49	0.06	0.21	-0.02
Gender	0.00	-0.01	-0.02	-0.04	-0.08	0.03	1.00	-0.03	-0.05	-0.04	-0.06	0.02	0.09	0.09	0.19
HR	-0.27	-0.18	-0.18	0.01	-0.04	0.13	-0.03	1.00	-0.02	0.03	-0.04	0.12	-0.05	0.22	0.08
MAP	-0.12	0.15	0.08	-0.04	-0.01	-0.06	-0.05	-0.02	1.00	-0.03	0.01	-0.01	0.06	-0.04	0.00
$PaCO_2$	-0.11	-0.39	0.44	-0.01	-0.14	0.18	-0.04	0.03	-0.03	1.00	-0.04	0.17	0.11	-0.05	-0.17
PaO_2	-0.06	0.00	0.02	0.03	0.00	0.11	-0.06	-0.04	0.01	-0.04	1.00	-0.09	-0.04	-0.09	0.03
PEEP	-0.22	-0.20	-0.05	0.01	-0.06	0.49	0.02	0.12	-0.01	0.17	-0.09	1.00	0.00	0.33	-0.15
Plat	0.00	0.05	0.15	-0.20	-0.12	0.06	0.09	-0.05	0.06	0.11	-0.04	0.00	1.00	-0.05	0.03
RR	-0.11	-0.21	-0.24	0.04	0.02	0.21	0.09	0.22	-0.04	-0.05	-0.09	0.33	-0.05	1.00	-0.31
V_T/PBW	0.03	0.07	-0.07	-0.01	0.00	-0.02	0.19	0.08	0.00	-0.17	0.03	-0.15	0.03	-0.31	1.00

Data are Pearson correlation coefficients.

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 4 - Absolute 60-day Mortality Difference Among Clusters per Trial and Model

FACTT trial (*n* = 998)			EDEN trial (*n* = 1000)		
Model	**Patients scored***	**Mortality difference among clusters**	**Model**	**Patients scored***	**Mortality difference among clusters**
6	93.5%	13.1%	7	77.7%	8.1%
2	57.4%	12.5%	8	77.7%	8.1%
5	65.5%	12.2%	6	84.1%	6.7%
8	70.2%	11.6%	5	71.7%	6.5%
7	70.2%	11.4%	9	84.7%	6.1%
1	57.4%	11.2%	3	77.7%	4.4%
4	70.2%	10.6%	4	77.7%	4.0%
9	93.5%	10.4%	2	57.7%	3.9%
3	70.2%	10.1%	10	87.3%	2.8%
10	98.8%	3.9%	1	57.7%	0.1%

* Number of patients without any missing data, allowing their assignment to one of the clusters.

eTable 5 - List of variables in each model assessed

Model	Demographics		Arterial Blood Gases			Laboratory Values				Vital Signs			Ventilator Variables		
	Age	Gender	pH	PaO_2	$PaCO_2$	Creat	Bili	HCO_3	Plat	MAP	RR	HR	FiO2	PEEP	V_T/PBW
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2			X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X		X	X	X	X		
4	X	X	X	X		X	X	X		X	X	X	X		
5			X	X	X	X	X	X	X	X	X	X	X		
6	X	X	X	X		X		X		X	X	X	X		
7			X	X	X	X	X	X		X	X	X	X		
8			X	X		X	X	X		X	X	X	X		
9			X	X	X			X		X	X	X			
10	X	X								X	X	X			

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 6 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	ALVEOLI			ARMA		
	Cluster 1 (*n* = 336)	**Cluster 2 (*n* = 157)**	***p* value**	**Cluster 1 (*n* = 279)**	**Cluster 2 (*n* = 100)**	***p* value**
Age, year*	53.0 (39.0 - 66.2)	46.0 (37.0 - 60.0)	0.007	49.0 (37.0 - 64.0)	47.5 (36.0 - 61.0)	0.180
Male gender - no. (%)	188 (56.0)	86 (54.8)	0.883	169 (60.6)	61 (61.0)	0.965
Body mass index, kg/m^2	27.0 (22.9 - 31.1)	25.2 (21.7 - 30.2)	0.050	25.8 (23.0 - 30.2)	24.4 (21.5 - 29.7)	0.057
Caucasian - no. (%)	263 (78.3)	102 (65.0)	0.002	220 (78.9)	65 (65.0)	0.009
Etiology - no. (%)			0.001			< 0.001
Pneumonia	130 (38.7)	66 (42.0)		83 (29.7)	30 (30.0)	
Sepsis	63 (18.8)	50 (31.8)		64 (22.9)	43 (43.0)	
Aspiration	55 (16.4)	19 (12.1)		44 (15.8)	14 (14.0)	
Trauma	33 (9.8)	5 (3.2)		43 (15.4)	4 (4.0)	
Other	55 (16.4)	17 (10.8)		45 (16.1)	9 (9.0)	
Prognostic scores						
APACHE III	71. (59.0 - 83.0)	93.0 (80.0 - 110.0)	< 0.001	77.0 (66.0 - 90.5)	97.0 (81.8 (110.0)	< 0.001
Use of vasopressor - no. (%)	65 (20.1)	80 (51.3)	< 0.001	77 (27.6)	52 (52.5)	< 0.001
Vital signs						
Temperature, ºC	37.6 (37.1 - 38.2)	37.7 (36.9 - 38.3)	0.778	37.6 (37.1 - 38.1)	37.6 (36.8 - 38.4)	0.803
Heart rate, bpm	97.5 (83.0 - 109)	111.0 (97.0 - 126)	< 0.001	101.0 (89.0 - 112.5)	118 (105.0 - 128.0)	< 0.001
Mean arterial Pressure, mmHg	77.3 (77.0 - 87.3)	73.3 (65.0 - 80.3)	< 0.001	78.0 (70.7 - 88.0)	70.5 (64.9 - 80.4)	< 0.001
SpO_2, %	96 (94 - 97)	95 (92 - 97)	0.005	95 (93 - 98)	95.5 (93 - 97)	0.799
Urine output in 24 hours, mL	2065 (1355 - 3255)	1433 (569 - 2189)	< 0.001	2100 (1375 - 3096)	1525 (816 - 2650)	0.001
Laboratory tests						
Hematocrit, %	31.0 (28.0 - 34.0)	31.0 (27.0 - 35.0)	0.617	30.0 (28.0 - 33.0)	31.0 (28.0 - 34.0)	0.299
White blood cell count, 10^9/L	11.7 (8.1 - 15.3)	10.7 (6.4 - 15.8)	0.166	11.9 (7.7 - 16.7)	9.8 (5.4 - 16.7)	0.057
Platelets, 10^9/L	173 (94 - 266)	141 (57 - 214)	0.001	139 (80 - 212)	125 (72 - 196)	0.260
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.5 (0.9 - 3.0)	< 0.001	1.0 (0.7 - 1.4)	1.8 (1.2 - 3.2)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.9 (0.4 - 1.8)	0.289	1.0 (0.6 - 2.1)	1.1 (0.7 - 2.7)	0.106
Arterial blood gas						
pH*	7.42 (7.38 - 7.45)	7.31 (7.24 - 7.36)	< 0.001	7.42 (7.38 - 7.47)	7.33 (7.28 - 7.37)	< 0.001
PaO_2, mmHg	78 (68 - 93)	74 (65 - 92)	0.082	75 (66 - 91)	81 (68 - 96)	0.106
PaO_2 / FiO_2	149 (109 - 192)	103 (74 - 136)	< 0.001	118 (83 - 160)	99 (68 - 137)	0.006

$PaCO_2$, mmHg	38 (34 - 43)	36 (31 - 42)	0.046	37 (31 - 41)	34 (28.8 - 39.2)	0.003
Bicarbonate, mmol/L	24 (21 - 27)	17 (13 - 20)	< 0.001	23 (20 - 26)	16 (13 - 19)	< 0.001
Ventilatory variables						
Tidal volume, mL	500 (437 - 600)	480 (400 - 572)	0.002	700 (600 - 750)	700 (550 - 700)	0.198
Per PBW, mL/kg PBW	8.0 (6.9 - 9.5)	7.4 (6.2 - 9.2)	0.006	10.1 (9.2 - 11.1)	10.6 (9.0 - 11.4)	0.383
Plateau pressure, cmH_2O	25.0 (21.0 - 30.0)	29.0 (24.0 - 33.0)	< 0.001	29.0 (24.0 - 34.0)	31.0 (27.0 - 36.0)	0.018
PEEP, cmH_2O	10 (5 - 10)	10 (8 - 14)	< 0.001	8 (5 - 10)	10 (5 - 12)	0.150
Respiratory rate, breaths/min	20 (15 - 25)	30 (24 - 35)	< 0.001	18 (14 - 21)	24 (18.8 - 28)	< 0.001
FiO_2	0.50 (0.44 - 0.65)	0.75 (0.60 - 1.00)	< 0.001	0.60 (0.50 - 0.70)	0.70 (0.59 - 0.96)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight.

eTable 7 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	SAILS			ART		
	Cluster 1 (*n* = 319)	**Cluster 2 (*n* = 188)**	***p* value**	**Cluster 1 (*n* = 211)**	**Cluster 2 (*n* = 298)**	***p* value**
Age, year*	57.0 (46.0 - 67.0)	53.5 (39.0 - 65.0)	0.035	54.0 (37.0 - 65.0)	50.0 (35.2 - 61.0)	0.075
Male gender - no. (%)	150 (47.0)	100 (53.2)	0.211	136 (64.5)	181 (60.7)	0.448
Body mass index, kg/m^2	28.5 (23.9 - 34.6)	29.8 (23.2 - 35.1)	0.903	28.8 (24.6 - 35.6)	28.4 (25.0 - 31.7)	0.367
Caucasian - no. (%)	250 (78.4)	140 (74.5)	0.369	---	---	---
Etiology - no. (%)			0.709			0.052
Pneumonia	228 (71.5)	127 (67.6)		113 (53.6)	171 (57.4)	
Sepsis	63 (19.7)	39 (20.7)		38 (18.0)	59 (19.8)	
Aspiration	19 (6.0)	15 (8.0)		13 (6.2)	16 (5.4)	
Trauma	3 (0.9)	1 (0.5)		10 (4.7)	2 (0.7)	
Other	6 (1.9)	6 (3.2)		37 (17.5)	50 (16.8)	
Prognostic scores				---	---	---
APACHE III	70.0 (56.0 - 84.0)	92.0 (75.0 - 105.8)	< 0.001			
SAPS III	---	---	---	62 (50 - 71)	66 (53 - 75)	0.010
Use of vasopressor - no. (%)	150 (47.8)	142 (78.5)	< 0.001	130 (61.6)	242 (81.2)	< 0.001
Vital signs						
Temperature, ºC	37.2 (36.7 - 37.8)	37.3 (36.7 - 38.0)	0.346	---	---	---
Heart rate, bpm	91.0 (80.5 - 103.0)	102.0 (88.8 - 117.0)	< 0.001	90.0 (73.0 - 103.0)	112.0 (97.2 - 126.0)	< 0.001
Mean arterial Pressure, mmHg	78.0 (69.5 - 88.0)	70.0 (63.0 - 78.)	< 0.001	80.0 (73.5 - 89.0)	75.0 (70.0 - 83.0)	< 0.001
SpO_2, %	96 (95 - 99)	96 (93 - 99)	0.270	---	---	---
Urine output in 24 hours, mL	1570 (852 - 2383)	920 (350 - 1665)	< 0.001	---	---	---
Laboratory tests						
Hematocrit, %	31 (27 - 35)	31 (28 - 37)	0.142	---	---	---
White blood cell count, $10^9/L$	13.6 (8.5 - 18.1)	15.4 (9.8 - 23.3)	0.009	---	---	---
Platelets, $10^9/L$	164 (96 - 238)	131 (80 - 223)	0.032	177 (120 - 292)	169 (90 - 256)	0.048
Creatinine, mg/dL	1.0 (0.7 - 1.5)	1.4 (0.9 - 2.6)	< 0.001	1.0 (0.7 - 1.5)	1.7 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.630	0.6 (0.4 - 1.2)	0.9 (0.4 - 1.7)	0.002
Arterial blood gas						
pH*	7.39 (7.35 - 7.44)	7.31 (7.24 - 7.35)	< 0.001	7.4 (7.3 - 7.4)	7.2 (7.2 - 7.3)	< 0.001
PaO_2, mmHg	82 (68 - 101)	86 (72 - 111.2)	0.112	118 (82 - 158)	104 (78 - 152)	0.065

PaO_2 / FiO_2	139 (98 - 195)	107 (74 - 159)	< 0.001	118 (82 - 158)	104 (78 - 152)	0.065
$PaCO_2$, mmHg	38 (34 - 45)	38 (32 - 44)	0.423	46 (41 - 56)	53 (42 - 65)	< 0.001
Bicarbonate, mmol/L	23 (20 - 26)	17 (14 - 21)	< 0.001	25.2 (22.5 - 28.8)	20.6 (17.8 - 23.4)	< 0.001
Ventilatory variables						
Tidal volume, mL	420 (360 - 480)	400 (340 - 450)	0.016	360 (320 - 400)	350 (300 - 397.8)	0.008
Per PBW, mL/kg PBW	6.4 (6.0 - 7.3)	6.1 (5.9 - 7.0)	0.030	6.0 (5.3 - 6.1)	5.9 (5.1 - 6.1)	0.034
Plateau pressure, cmH_2O	22.0 (18.0 - 27.0)	25.0 (20.0 - 29.0)	0.003	24.0 (21.0 - 28.0)	27.0 (23.0 - 30.0)	< 0.001
PEEP, cmH_2O	8 (5 - 10)	10 (8 - 13)	0.001	10 (10 - 14)	12 (10 - 14)	< 0.001
Respiratory rate, breaths/min	23 (19 - 27)	30 (24 - 35)	< 0.001	24 (20 - 28)	30 (24 - 34)	< 0.001
FiO_2	0.50 (0.40 - 0.60)	0.70 (0.50 - 0.90)	< 0.001	0.70 (0.60 - 0.80)	0.80 (0.70 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight...

eTable 8 - Biomarker levels by study and cluster

	ARMA				ALVEOLI			
	Subphenotype A (*n* = 279)	**Subphenotype B (*n* = 100)**	**Median Difference (95% CI)**	***p* value**	**Subphenotype A (*n* = 336)**	**Subphenotype B (*n* = 157)**	**Median Difference (95% CI)**	***p* value**
ICAM-1	654.0 (399.0 - 959.4)	888.0 (550.0 - 1365.3)	234 (60.3 to 407.8)	0.002	847.9 (585.7 - 1227.1)	1070.4 (748.2 - 1588.8)	219.4 (90.4 to 348.4)	< 0.001
IL-6	214.0 (91.8 - 553.5)	966.0 (291.0 - 2200.0)	749.1 (589.9 to 908.2)	< 0.001	182.5 (85.5 - 435.2)	775.0 (148.0 - 2846.5)	592 (515.5 to 668.6)	< 0.001
PAI-1	65.3 (37.8 - 109.5)	101.7 (50.8 - 291.6)	41 (18.3 to 63.7)	0.001	Not assessed	Not assessed	---	---
IL-8	46.0 (2.0 - 91.0)	106.9 (43.8 - 281.4)	60.9 (35.6 to 86.2)	< 0.001	Not assessed	Not assessed	---	---
IL-10	16.0 (0.0 - 40.3)	47.9 (0.0 - 120.7)	31.9 (20.2 to 43.6)	< 0.001	Not assessed	Not assessed	---	---
TNFR-I	2604.0 (1950.0 - 3777.0)	6897.0 (3622.5 - 12281.5)	4293 (3323.6 to 5262.4)	< 0.001	Not assessed	Not assessed	---	---
TNFR-II	6581.0 (4958.0 - 9658.0)	18611.0 (12262.5 - 35652.0)	12030 (9577.5 to 14482.5)	< 0.001	Not assessed	Not assessed	---	---
SPA	29.0 (11.8 - 68.0)	25.0 (10.5 - 40.0)	-4 (-19.9 to 11.9)	0.398	Not assessed	Not assessed	---	---
SPD	76.0 (36.2 - 145.2)	59.0 (30.0 - 125.0)	-18 (-52.6 to 16.6)	0.254	Not assessed	Not assessed	---	---
VW	308.0 (165.5 - 431.0)	384.0 (246.0 - 549.0)	76 (-26.5 to 178.5)	0.045	Not assessed	Not assessed	---	---

Data are median (quartile 25th - quartile 75th).

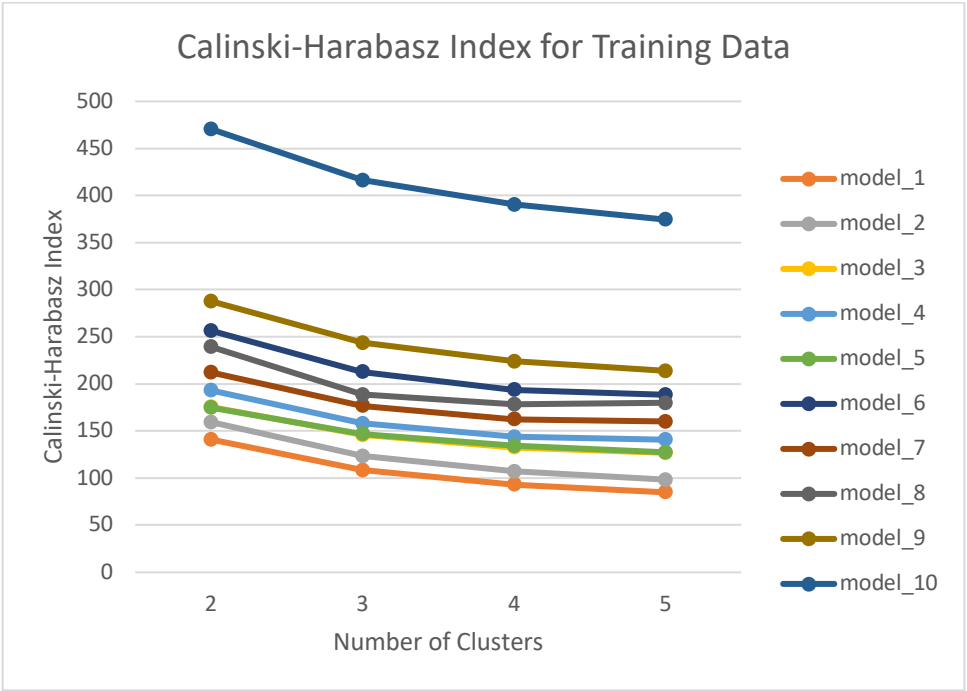
Abbreviations: 95%CI denotes 95% confidence interval, ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

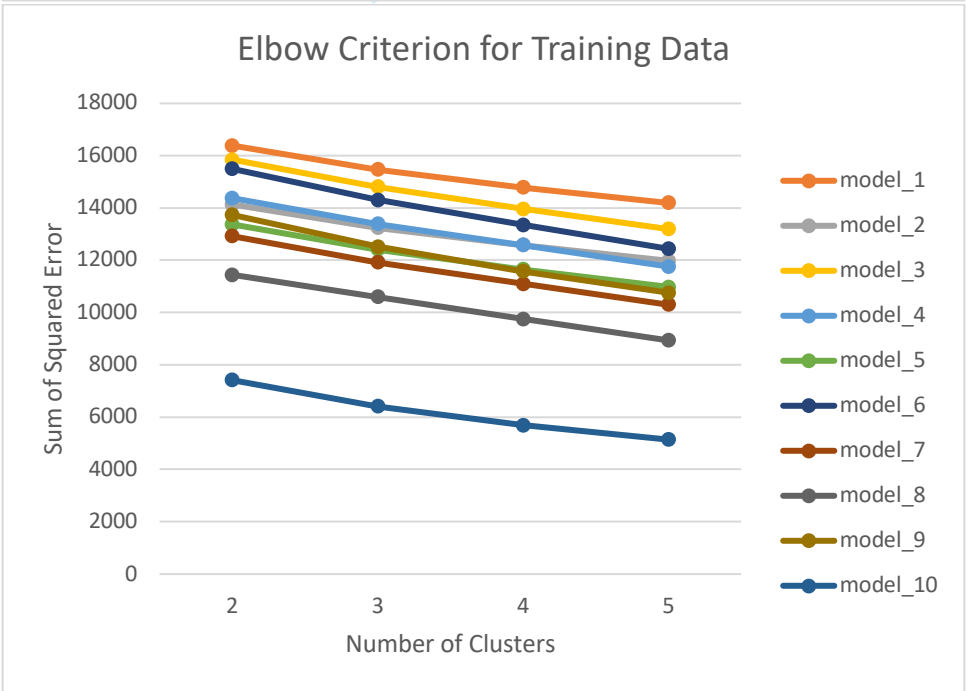
eTable 9 - Percentage of missingness in biomarker levels measured on day of randomization, on ARMA and ALVEOLI trials for patients with an assigned subphenotype

Biomarker	ARMA (*n* = 379)		ALVEOLI (*n* = 493)	
	Subphenotype A	Subphenotype B	Subphenotype A	Subphenotype B
ICAM-1	43%	31%	4%	3%
IL-6	41%	33%	4%	4%
PAI-1	42%	32%	Not assessed	Not assessed
IL-8	41%	33%	Not assessed	Not assessed
IL-10	42%	33%	Not assessed	Not assessed
TNFR-I	68%	61%	Not assessed	Not assessed
TNFR-II	68%	61%	Not assessed	Not assessed
SPA	67%	61%	Not assessed	Not assessed
SPD	67%	61%	Not assessed	Not assessed
VW	67%	61%	Not assessed	Not assessed

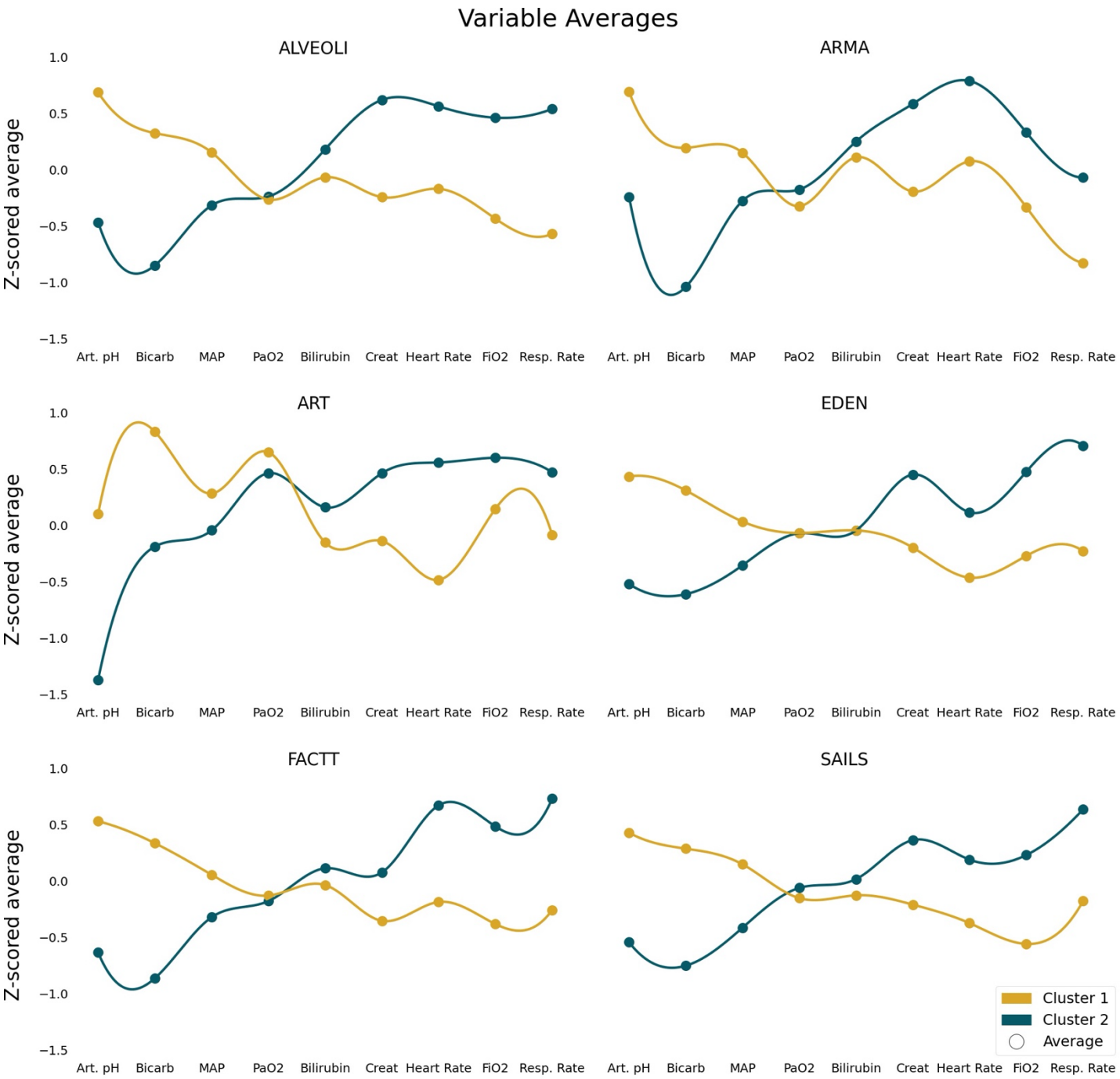
Abbreviations: ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

eFigure 1 - Calinski-Harabasz Index and Elbow Method for Each of the 10 Models





eFigure 2 - Variable Averages for Each Study



The circles represent the averages for each variable. The colored lines are exclusively to help visualize the opposite trends of the variables on the different clusters.

Abbreviations: Art. pH is arterial pH, Bicarb is bicarbonate, MAP is mean arterial pressure, Creat is creatinine and Resp. Rate is respiratory rate

Reference:

1. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW. Reappraisal of Ventilator-Free Days in Critical Care Research. Am J Respir Crit Care Med.2019 Oct 1;200(7):828-836. doi: 10.1164/rccm.201810-2050CP.

Reporting checklist for prediction model development/validation.

Based on the TRIPOD guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the TRIPODreporting guidelines, and cite them as:

Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement.

		Reporting Item	Page Number
Title			
	#1	Identify the study as developing and / or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract			
	#2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
	#3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	6
	#3b	Specify the objectives, including whether the study describes the	6

		development or validation of the model or both.	
Methods			
Source of data	#4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8
Source of data	#4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	8
Participants	#5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8
Participants	#5b	Describe eligibility criteria for participants.	8
Participants	#5c	Give details of treatments received, if relevant	8
Outcome	#6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	9
Outcome	#6b	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	#7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured	8
Predictors	#7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	#8	Explain how the study size was arrived at.	8
Missing data	#9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	9
Statistical analysis methods	#10a	If you are developing a prediction model describe how predictors were handled in the analyses.	N/A
Statistical analysis methods	#10b	If you are developing a prediction model, specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	N/A
Statistical analysis methods	#10c	If you are validating a prediction model, describe how the predictions were calculated.	N/A
Statistical analysis	#10d	Specify all measures used to assess model performance and, if relevant,	10

methods		to compare multiple models.	
Statistical analysis methods	#10e	If you are validating a prediction model, describe any model updating (e.g., recalibration) arising from the validation, if done	N/A
Risk groups	#11	Provide details on how risk groups were created, if done.	11
Development vs. validation	#12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	10
Results			
Participants	#13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	12
Participants	#13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	12
Participants	#13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	12
Model development	#14a	If developing a model, specify the number of participants and outcome events in each analysis.	12
Model development	#14b	If developing a model, report the unadjusted association, if calculated between each candidate predictor and outcome.	N/A
Model specification	#15a	If developing a model, present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
Model specification	#15b	If developing a prediction model, explain how to the use it.	N/A
Model performance	#16	Report performance measures (with CIs) for the prediction model.	14
Model-updating	#17	If validating a model, report the results from any model updating, if done (i.e., model specification, model performance).	N/A
Discussion			
Limitations	#18	Discuss any limitations of the study (such as nonrepresentative sample,	19

		few events per predictor, missing data).	
Interpretation	#19a	For validation, discuss the results with reference to performance in the development data, and any other validation data	17
Interpretation	#19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	17
Implications	#20	Discuss the potential clinical use of the model and implications for future research	20
Other information			
Supplementary information	#21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	22
Funding	#22	Give the source of funding and the role of the funders for the present study.	22

This checklist was completed on 07. May 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai

BMJ Open

Identification of Acute Respiratory Distress Syndrome subphenotypes denovo using routine clinical data: a retrospective analysis of ARDS clinical trials

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-053297.R2
Article Type:	Original research
Date Submitted by the Author:	29-Nov-2021
Complete List of Authors:	Duggal, Abhijit ; Cleveland Clinic, Department of Critical Care Medicine Kast, Rachel; Endpoint Health Inc, Department of Clinical Data Science Van Ark, Emily; Endpoint Health Inc, Department of Clinical Data Science Bulgarelli, Lucas; Endpoint Health Inc, Department of Clinical Data Science Siuba, Matthew T.; Cleveland Clinic, Department of Critical Care Medicine Osborn, Jeff; Endpoint Health Inc, Department of Clinical Data Science Rey, Diego; Endpoint Health Inc, Department of Clinical Data Science Zampieri, Fernando; HCor Research Institute Cavalcanti, Alexandre ; HCor Research Institute Maia, Israel; Hospital do Coracao Paisani, Denise M; HCor Research Institute Laranjeira, Ligia N; HCor Research Institute Serpa Neto, Ary; Monash University, Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine,; Hospital Israelita Albert Einstein, Critical Care Medicine Deliberato, Rodrigo Octávio; Endpoint Health Inc, Department of Clinical Data Science
Primary Subject Heading:	Intensive care
Secondary Subject Heading:	Respiratory medicine
Keywords:	Adult intensive & critical care < INTENSIVE & CRITICAL CARE, Adult intensive & critical care < ANAESTHETICS, RESPIRATORY MEDICINE (see Thoracic Medicine)

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Identification of Acute Respiratory Distress Syndrome subphenotypes de novo using routine clinical data: a retrospective analysis of ARDS clinical trials

Abhijit Duggal MD[1‡], Rachel Kast PhD[2‡], Emily Van Ark PhD[2‡], Lucas Bulgarelli BSc[2‡], Matthew T. Siuba DO[1], Jeff Osborn[2], Diego Rey PhD[2], Fernando G Zampieri MD PhD[3], Alexandre B Cavalcanti MD PhD[3], Israel S Maia MD[3], Denise M Paisani PhD[3], Ligia N Laranjeira[3], Ary Serpa Neto MD MSc PhD,[4,5,6,7,8] Rodrigo Octávio Deliberato MD PhD[2]

‡ Authors contributed equally

1. Department of Critical Care Medicine, Respiratory Institute, Cleveland Clinic, Cleveland, Ohio, USA.
2. Department of Clinical Data Science, Endpoint Health Inc, Palo Alto, California, USA.
3. HCor Research Institute, São Paulo, Brazil
4. Australian and New Zealand Intensive Care Research Centre (ANZIC-RC), School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia.
5. Department of Critical Care, Melbourne Medical School, University of Melbourne, Austin Hospital, Melbourne, Australia.
6. Department of Intensive Care, Austin Hospital, Melbourne, Australia.
7. Data Analytics Research and Evaluation (DARE) Centre, Austin Hospital, Melbourne, Australia.
8. Department of Critical Care Medicine, Hospital Israelita Albert Einstein, São Paulo, Brazil.

Correspondence:
Abhijit Duggal MD
Address: 9500 Euclid Ave, L2-330, Cleveland, Ohio, 44195
E-mail: duggala2@ccf.org

Word count (Abstract): 236 words

Word count (Text): 2814 words

Number of figures: 2 figures

Number of tables: 3 tables

Supplementary Material: 01

Key words: Subphenotype, machine learning, ARDS, critical care, clinical data, clustering

ABSTRACT

Objectives: The acute respiratory distress syndrome (ARDS) is a heterogeneous condition, and identification of subphenotypes may help in better risk stratification. Our study objective is to identify ARDS subphenotypes using new simpler methodology and readily available clinical variables.

Setting: This is a retrospective Cohort Study of ARDS trials. Data from the U.S. ARDSNet trials and from the international ART trial.

Participants: 3763 patients from ARDSNet datasets and 1010 patients from the ART dataset.

Primary and secondary outcome measures: The primary outcome was 60-day or 28-day mortality, depending on what was reported in the original trial. K-means cluster analysis was performed to identify subgroups. Sets of candidate variables were tested to assess their ability to produce different probabilities for mortality in each cluster. Clusters were compared to biomarker data, allowing identification of subphenotypes.

Results: Data from 4,773 patients was analyzed. Two subphenotypes (A and B) resulted in optimal separation in the final model, which included nine routinely collected clinical variables, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. Participants in subphenotype B showed increased levels of pro-inflammatory markers, had consistently higher mortality, lower number of ventilator-free days at day 28, and longer duration of ventilation compared to patients in the subphenotype A.

Conclusions: Routinely available clinical data can successfully identify two distinct subphenotypes in adult ARDS patients. This work may facilitate implementation of precision therapy in ARDS clinical trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Largest cohort of patients used to identify subphenotypes of ARDS patients.
- Subphenotypes were validated in the population of a large international ARDS randomized controlled trial.
- Subphenotypes were identified by using only routinely collected clinical data.
- Our use of data exclusively from randomized controlled trials does not prove generalizability to unselected ARDS populations.
- The clinical utility of the subphenotypes has to be validated in a prospective study.

INTRODUCTION

The Berlin definition of acute respiratory distress syndrome (ARDS) encompasses acute hypoxemic respiratory failure due to a wide variety of etiologies [1]. Due to this inclusion of heterogeneous conditions within the syndrome, there are significant clinical and biological differences that make ARDS challenging to treat [2,3]. These differences amongst ARDS patients are associated with variation in risk of disease development and progression [3,4], potentially generating differential responses to treatments and interventions [5–10]. Despite evidence, clinical risk stratification of ARDS patients still solely depends on PaO_2/FiO_2 ratios [11,12], possibly misleading the interpretation of results in clinical trials and clinicians when evaluating treatment options for patients [13].

Therefore, identifying groups of patients who have similar clinical, physiologic, or biomarker traits becomes relevant [6,14] as it can help with stratification of patients producing better targeted therapies and interventions [15]. These different groups can be defined as ARDS subphenotypes [4,14]. Two ARDS subphenotypes have been consistently identified in previous studies [6–10,16–18]. However, these models are complex, and significant barriers exist in their implementation and use in clinical practice. Existing models use up to 40 predictor variables, including biomarkers and other variables that are not readily available at the bedside [6–10,16–18]. These limitations explain the current status quo of ARDS care, where clinicians must depend on the limited prognostic value of PaO_2/FiO_2 ratios instead of biologically distinct subphenotypes.

We hypothesized that the use of a simpler methodology and a small number of easily available clinical variables could identify new ARDS subphenotypes and thus provide the means to allow future implementation of bedside stratification.

METHODS

Data source and participants

We performed a retrospective study using a de-identified dataset pooling data from six randomized clinical trials in patients with ARDS, namely: ARMA, ALVEOLI, FACTT, EDEN, SAILS, and ART [19–24]. Patients in ARMA, ALVEOLI, FACTT, EDEN, and SAILS trials were eligible if they met the American-European consensus for ARDS, including patients with a PaO_2 / FiO_2 ratio < 300 up to 48 hours before enrollment. From 1996 to 2013, these trials enrolled 902, 549, 1000, 1000, and 745 patients, respectively, and tested a variety of interventions [19–23]. Between 2011 and 2017 the international ART study enrolled 1010 adult patients diagnosed with moderate to severe ARDS according to the Berlin definition (PaO_2 / FiO_2 ratio < 200) for less than 72 hours of duration and assessed two different ventilatory strategies [24]. To avoid biases due to high mortality in the high tidal volume group of the ARMA study [19], which has not been standard of care since the beginning of 2000, only 473 patients receiving low tidal volume in that study were included.

Predictors

Six clinical trials were assessed to identify a set of clinical variables recorded closest to time of randomization which were most commonly available across all datasets. The list of potential candidates was then further refined to include only those that are frequently observed in the routine care of ARDS patients at the time of its diagnosis according to judgement provided by ICU physicians who participated in this study. To develop a clustering algorithm for potential rapid translation into clinical use, elements which would not be commonly found in the electronic health records (EHR) at the time of ARDS

diagnosis, such as biomarker levels, ARDS risk factors, organ support apart from mechanical ventilation settings, and severity scores, were excluded from model development. The treatment assignment in the original trials, and clinical outcomes were not considered in the model development.

After all assessment, 16 variables that are routinely collected as part of the usual care and which were uniformly present in all the trials were considered, including: age, gender, arterial pH, PaO_2, $PaCO_2$, bicarbonate, creatinine, bilirubin, platelets, heart rate, respiratory rate, mean arterial pressure, positive end-expiratory pressure (PEEP), plateau pressure, FiO_2, and tidal volume adjusted for predicted body weight (mL/kg PBW). The PBW was calculated as equal to 50 + 0.91 (centimeters of height – 152.4) in males, and 45.5 + 0.91 (centimeters of height – 152.4) in females [18]. These variables were grouped into five domains named demographics, arterial blood gases, laboratory values, vital signs, and ventilatory variables. Plateau pressure was excluded due to a high rate of missingness across the trials included in the training set. Amount of missing data in the training datasets is reported in **eTable 1**.

Outcomes

The primary outcome was 60-day mortality for all ARDSnet trials, and 28-day mortality for ART trial. Secondary outcomes included 90-day mortality, number of ventilator free days at day 28 [25], and the duration of mechanical ventilation in survivors within the first 28 days post enrollment.

Data preparation

Data preprocessing was performed before modeling, and the pooled dataset was assessed for completeness and consistency. Patients with values out of the plausible

physiological range for a specific variable were excluded from the final analysis (described in **eTable 2**). The training dataset was constructed using data from the two largest ARDSnet trials, EDEN and FACTT. The validation dataset was sourced from the four remaining trials: ALVEOLI, ARMA, SAILS, and ART. Means and standard deviations for z-scoring variables were calculated from the training dataset and subsequently applied to the validation data.

Statistical analysis

Baseline and outcome data were presented according to the assigned cluster. Continuous variables were presented as medians with their interquartile ranges and categorical variables as total number and percentage. Proportions were compared using Fisher exact tests and continuous variables were compared using the Wilcoxon rank-sum test. Study outcomes were further compared using the median and mean absolute differences for continuous and categorical values, respectively.

Model development and validation

For the model development, the K-means clustering algorithm was used. K-means is one of the simplest and most used classes of clustering algorithms. In critical care research, unsupervised machine learning techniques have already been used in several studies, attempting to find homogeneous subgroups within a broad heterogeneous population [26]. This specific algorithm identifies a K number of clusters in a dataset by finding K centroids within the n-dimensional space of clinical features [26].

For feature selection, different sets of candidate variables were tested to assess their ability to produce significantly different mortality probabilities in each cluster using the minimum amount of readily available clinical data. For each set of candidate variables,

the optimal number of clusters was determined by comparing models with between 2 and 5 clusters, using the Elbow method [27] and the Calinski-Harabasz index [28]. Information about the methods for selecting the number of clusters are provided in the supplemental material.

The following steps were performed for the final model selection: 1) all predictors were assessed for correlation (**eTable 3**); and 2) ten different combinations of the proposed variables were investigated. These combinations were developed based on the perceived clinical importance of each variable and its combinations. All 10 models were tested for the optimal number of clusters based on both the Elbow method and the Calinski-Harabasz index, as described above. The models were then compared, aiming for the minimum set of variables with high 60-day mortality separation. The description of each model is shown in **eTable 4**.

Biological and clinical characteristics of the clusters were evaluated using clinical, laboratory, and (when available) biomarker data to establish subphenotypes [4]. All iterations in model development were done on the training set and the generalizability of the final model was assessed using the validation dataset. K-means clustering analysis is structured to ignore cases with missing data. No assumption was made for missingness, and we therefore conducted a complete case analysis. Model development and evaluation was performed using Python version 3.8 and scikit-learn 0.23.1.

Patient and public involvement

There was no patient involvement in this study.

Data availability

Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI and SAILS) is publicly available from the NHLBI ARDS Network and data from the ART trial can be requested from study authors.

RESULTS

Participants

Data from 4777 clinical trial patients were considered for inclusion. In total, 4 patients were excluded for having clinical measurements outside plausible range. The remaining 1998 patients from EDEN and FACTT trials were included in the training set, while the 2775 patients from ARMA, ALVEOLI, SAILS, and ART were included in the validation cohort.

Baseline characteristics of the patients in the training and validation sets are presented in **Table 1**. Pneumonia was the prevailing etiology followed by sepsis and aspiration in all trials. Between 29.3% to 72.7% of the patients were receiving vasopressors at the time of randomization. At randomization, PaO_2 / FiO_2 ratio ranged from 112 (75 - 158) to 134 (96 - 185) mmHg, and PEEP from 8 (5 - 10) to 12 (10 - 14) cmH_2O across trials. Mortality at 60 days for the ARDSnet trials ranged from 22.7% to 30.1%, while in the ART trial mortality at 28 days was 58.8%.

Table 1 - Baseline Characteristics and Clinical Outcomes in the Included Trials

	Training set (n = 1998)		Validation set (n = 2775)			
	EDEN (n = 1000)	**FACTT (n = 998)**	**ALVEOLI (n = 549)**	**ARMA (n = 472)**	**ART (n = 1010)**	**SAILS (n = 744)**
Age, year	52.0 (42.0 - 63.0)	49.0 (38.0 - 60.8)	50.0 (39.0 - 65.0)	50.0 (37.8 - 65.0)	52.0 (36.0 - 64.0)	55.0 (42.0 - 66.0)
Male gender - no. (%)	510 (51.0)	533 (53.4)	302 (55.0)	285 (60.4)	631 (62.5)	365 (49.0)
Etiology - no. (%)						
Pneumonia	650 (65.0)	471 (47.2)	221 (40.3)	145 (30.7)	555 (55.0)	526 (70.7)
Sepsis	147 (14.7)	231 (23.1)	120 (21.9)	125 (26.5)	196 (19.4)	147 (19.8)
Aspiration	96 (9.6)	149 (14.9)	84 (15.3)	72 (15.3)	58 (5.7)	49 (6.6)
Trauma	36 (3.6)	74 (7.4)	45 (8.2)	59 (12.5)	31 (3.1)	6 (0.8)
Other	71 (7.1)	73 (7.3)	79 (14.4)	71 (15.0)	170 (16.8)	16 (2.2)
Severity of Illness*	73.0 (59.0 - 89.0)	78.0 (62.0 - 94.0)	78.0 (64.0 - 93.0)	83.0 (70.0 - 97.0)	63.0 (50.2 – 75.0)	76.0 (61.0 - 92.0)
Vasopressors - no. (%)	489 (48.9)	397 (40.5)	156 (29.3)	147 (31.3)	734 (72.7)	395 (54.2)
Laboratory tests						

White blood cell count, 109/L	12.0 (7.8 - 16.7)	11.8 (7.2 - 17.1)	11.6 (7.7 - 15.7)	11.5 (7.5 - 16.2)	---	13.9 (8.7 - 20.0)
Platelets, 109/L	169 (108 - 241)	183 (106 - 258)	157 (83 - 247)	135 (80 - 211)	175 (106 - 263)	167 (96 - 247)
Creatinine, mg/dL	1.2 (0.8 - 2.0)	1.0 (0.7 - 1.5)	1.0 (0.7 - 1.7)	1.1 (0.8 - 1.7)	1.3 (0.8 - 2.2)	1.0 (0.7 - 1.7)
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.8 (0.5 - 1.5)	1.0 (0.6 - 2.1)	0.8 (0.4 - 1.5)	0.8 (0.5 - 1.4)
Arterial blood gas						
pH*	7.36 (7.30 - 7.42)	7.37 (7.30 - 7.43)	7.40 (7.34 - 7.44)	7.41 (7.35 - 7.45)	7.28 (7.19 - 7.36)	7.37 (7.31 - 7.42)
PaO2, mmHg	83 (68 - 108)	79 (67 - 100)	77 (67 - 93)	76.5 (67 - 93)	112 (81 - 155)	83 (69 - 103)
PaO2 / FiO2	125 (86 - 178)	118 (80 - 163)	134 (96 - 185)	112 (75 - 158)	112 (81 - 155)	133 (89 - 178)
PaCO2, mmHg	38 (34 - 45)	39 (34 - 45)	38 (33 - 43)	36 (31 - 41)	50 (42 - 62)	39 (34 - 45)
Bicarbonate, mmol/L	21.0 (18.0 - 25.0)	21.0 (17.4 - 25.0)	22.0 (18.0 - 26.0)	22.0 (18.0 - 25.0)	22.9 (19.4 - 26.3)	22.0 (18.0 - 25.0)
Ventilatory variables						
Tidal volume, mL	410 (360 - 470)	450 (400 - 510)	500 (420 - 600)	700 (600 - 750)	350 (308 - 400)	400 (350 - 460)
Per PBW, mL/kg PBW	6.3 (6.0 - 7.3)	7.1 (6.1 - 8.1)	7.9 (6.6 - 9.4)	10.2 (9.0 - 11.3)	5.9 (5.1 - 6.1)	6.2 (6.0 - 7.1)
Plateau pressure, cmH2O	24.0 (20.0 - 27.0)	26.0 (22.0 - 30.0)	26.0 (22.0 - 31.0)	29.0 (24.8 - 34.0)	26.0 (22.0 - 29.0)	24.0 (19.0 - 28.0)
PEEP, cmH2O	10 (5 - 12)	10 (5 - 12)	10 (5 - 12)	8 (5 - 10)	12 (10 - 14)	10 (5 - 11)
FiO2	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.80)	0.60 (0.50 - 0.74)	0.70 (0.60 - 1.00)	0.60 (0.40 - 0.70)
Clinical outcomes						
28-day mort. - no. (%)	---	---	---	---	594 (58.8)	---
60-day mort. - no. (%)	227 (22.7)	268 (26.9)	144 (26.2)	141 (30.1)	---	199 (26.7)
90-day mort. - no. (%)	233 (23.3)	283 (28.6)	148 (27.5)	143 (30.8)	---	204 (27.4)
Ventilator-free days, day 28	20.0 (0.0 - 24.0)	17.0 (0.0 - 23.0)	18.0 (0.0 - 24.0)	13.0 (0.0 - 23.0)	0.0 (0.0 - 13.0)	20.0 (0.0 - 25.0)
Ventilator days in survivors	7.0 (4.0 - 13.0)	8.0 (5.0 - 16.0)	8.0 (4.0 - 14.0)	8.0 (4.0 - 15.0)	13.0 (8.0 - 20.0)	6.0 (4.0 - 11.0)

Data are median (quartile 25th - quartile 75th) or N (%)
Abbreviations: 28-day mort. is 28-day mortality, 60-day mort. is 60 days mortality, and 90-day mort. is 90-day mortality.
* Except for ART, that uses SAPS-3, all studies use APACHE-IV

Predictor variables and model selection

The correlation between the 15 variables selected for clustering is shown in **eTable 3**. The strongest correlation was between PEEP and FiO_2 ($r = 0.49$). The comparison of the 10 models regarding the optimal number of clusters based on both the Elbow method and the Calinski-Harabasz index is shown in **eFigure 1**. In all models and methods, two clusters were a better fit than a higher number of clusters.

Across the ten models, absolute mortality difference between cluster 1 and cluster 2 ranged from 3.9% to 13.1% for the FACTT study and between 0.1% to 8.1% for EDEN (**eTable 4**). The models with the highest 60-day absolute mortality separation between the clusters for each of the two trials in the training set were then further evaluated. Models 6, 5, and 8 were consistently amongst the models with highest separation (**eTable 4**). Model 8 was selected for further investigation, as it had the fewest variables (**eTable 5**).

Clinical characteristics of each cluster

Based on model 8, only nine clinical and laboratory variables were needed to identify the two distinct clusters in ARDS patients, namely: heart rate, mean arterial pressure, respiratory rate, bilirubin, bicarbonate, creatinine, PaO_2, arterial pH, and FiO_2. For each variable in the model, opposing measurements could be observed for each cluster (**Figure 1** and **eFigure 2**). For the ARDSnet trials, the incidence of cluster 1 patients varied from 57.8% (EDEN) to 73.6% (ARMA), and 41.5% of ART patients were part of cluster 1. Across all trials, patients in cluster 2 had higher severity of illness, rate of vasopressor, heart rate, respiratory rate, creatinine, and bilirubin, as well as lower platelets, pH, BUN, and bicarbonate compared to patients in cluster 1 (**Table 2, eTables 6** and **7**). In addition, 28-, 60-, and 90-day mortality rate was higher in patients in cluster 2 in all trials (**Table 3**). Likewise, for each trial, the number of ventilator-free days at day 28 was lower in patients in cluster 2 compared to cluster 1, and duration of ventilation in survivors was longer in cluster 1.

Table 2 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Trials in the Training Set						
	FACTT			EDEN		
	Cluster 1 (n = 407)	Cluster 2 (n = 294)	p value	Cluster 1 (n = 449)	Cluster 2 (n = 328)	p value

Age, year*	50.0 (40.0 - 63.0)	47.0 (36.0 - 58.0)	0.002	53.0 (44.0 - 63.0)	51.0 (41.0 - 62.2)	0.183
Male gender - no. (%)	223 (54.8)	151 (51.4)	0.411	233 (51.9)	168 (51.2)	0.910
Body mass index, kg/m2	27.5 (23.3 - 32.1)	27.4 (23.0 - 32.7)	0.938	29.1 (24.6 - 34.5)	28.5 (23.4 - 35.1)	0.476
Caucasian - no. (%)	269 (66.1)	177 (60.2)	0.129	349 (81.5)	237 (75.7)	0.067
Etiology - no. (%)			< 0.001			0.003
Pneumonia	201 (49.4)	139 (47.3)		296 (65.9)	217 (66.2)	
Sepsis	78 (19.2)	101 (34.4)		50 (11.1)	60 (18.3)	
Aspiration	67 (16.5)	30 (10.2)		45 (10.0)	27 (8.2)	
Trauma	24 (5.9)	8 (2.7)		24 (5.3)	5 (1.5)	
Other	37 (9.1)	16 (5.4)		34 (7.6)	19 (5.8)	
Prognostic scores						
APACHE III	69.0 (56.0 - 84.0)	91 (76.0 - 105.0)	< 0.001	66.0 (54.0 - 79.0)	84.0 (71.0 - 100.2)	< 0.001
Use of vasopressor - no. (%)	118 (29.5)	189 (64.9)	< 0.001	187 (41.6)	209 (63.7)	< 0.001
Vital signs						
Temperature, ºC	37.5 (36.8 - 38.2)	37.6 (37.0 - 38.4)	0.371	37.3 (36.8 - 37.8)	37.3 (36.7 - 38.1)	0.212
Heart rate, bpm	95.0 (81.0 - 110.0)	114 (102 - 126)	< 0.001	89 (77 - 102)	101 (89 - 116)	< 0.001
Mean arterial Pressure, mmHg	76.0 (68.0 - 88.0)	71.0 (65.0 - 80.8)	< 0.001	77.0 (68.0 - 84.0)	71.0 (66.0 - 80.0)	< 0.001
SpO2, %	96 (93 - 98)	95 (92 - 97)	< 0.001	96 (94 - 98)	95 (92 - 98)	0.032
Urine output in 24 hours, mL	1785 (1192 - 2853)	1370 (842 - 2446)	< 0.001	1505 (977 - 2250)	1165 (566 - 1816)	< 0.001
Laboratory tests						
Hematocrit, %	30.0 (26.0 - 33.0)	30.0 (24.2 - 35.0)	0.272	30.0 (26.0 - 34.0)	30.0 (26.0 - 35.0)	0.919
White blood cell count, 109/L	11.6 (7.3 - 16.3)	11.7 (5.6 - 17.9)	0.972	11.4 (7.7 - 15.5)	12.7 (7.7 - 19.0)	0.019
Platelets, 109/L	195 (118.5 - 268)	158 (87 - 237)	< 0.001	163 (108 - 241)	164 (103 - 227)	0.552
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.4 (1.0 - 2.0)	< 0.001	1.0 (0.7 - 1.5)	1.6 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.7 (0.5 - 1.3)	0.9 (0.5 - 2.0)	0.003	0.8 (0.5 - 1.3)	0.8 (0.5 - 1.7)	0.128
Arterial blood gas						
pH*	7.41 (7.36 - 7.45)	7.29 (7.23 - 7.35)	< 0.001	7.40 (7.35 - 7.44)	7.30 (7.24 - 7.35)	< 0.001
PaO2, mmHg	78 (68 - 100)	78 (65 - 99)	0.240	83 (70 - 107)	81 (67 - 107)	0.416
PaO2 / FiO2	132 (92 - 173)	89 (65 - 126)	< 0.001	133 (98 - 193)	101 (73 - 162)	< 0.001
PaCO2, mmHg	39 (34 - 44)	38.5 (33 - 47.9)	0.877	38 (34 - 44)	38 (33 - 46)	0.55
Bicarbonate, mmol/L	24.0 (21.0 - 27.0)	17.0 (14.0 - 20.0)	< 0.001	23.0 (21.0 - 26.0)	18.5 (15.0 - 21.0)	< 0.001
Ventilatory variables						
Tidal volume, mL	450 (400 - 530)	450 (382 - 500)	0.009	420 (356 - 487)	400 (350 - 450)	0.032
Per PBW, mL/kg PBW	7.1 (6.3 - 8.4)	7.0 (6.0, 8.0)	0.058	6.3 (6.0 - 7.5)	6.1 (6.0 - 7.3)	0.079
Plateau pressure, cmH2O	25.0 (20.0 - 29.0)	28.0 (24.0 - 32.0)	< 0.001	23.0 (19.0 - 27.0)	24.0 (21.0 - 28.0)	0.004
PEEP, cmH2O	8 (5 - 10)	10 (8 - 14)	< 0.001	10 (5 - 10)	10 (8 - 14)	< 0.001
Respiratory rate, breaths/min	22 (18 - 27)	30 (24 - 35)	< 0.001	22 (19 - 26)	30 (25 - 35)	< 0.001
FiO2	0.50 (0.40 - 0.70)	0.80 (0.60 - 1.00)	< 0.001	0.60 (0.45 - 0.70)	0.80 (0.60 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25th - quartile 75th) or N (%)
Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, VT/PBW denotes tidal volume per predicted body weight.

Table 3 - Clinical Outcomes According to Clusters in Each Trial				
	Cluster 1	**Cluster 2**	**Difference (95% CI)**	**p value**
Training set				
FACTT	n = 407	n = 294		
60-day mortality - no. (%)	94 (23.1)	102 (34.7)	11.6% (4.9% to 18.3%)	0.001
90-day mortality - no. (%)	103 (25.4)	106 (36.3)	10.9% (4.1% to 17.8%)	0.002
Ventilator-free days at day 28	19.0 (0.0 - 24.0)	10.0 (0.0 - 21.0)	-9.0 (-11.9 to -6.1)	< 0.001
Duration of ventilation in survivors, days	8.0 (4.0 - 13.0)	10.0 (7.0 - 19.0)	2.0 (0.5 to 3.5)	< 0.001
EDEN	n = 449	n = 328		
60-day mortality - no. (%)	87 (19.4)	90 (27.4)	8.1% (2.1% to 14.0%)	0.010
90-day mortality - no. (%)	90 (20.0)	93 (28.4)	8.3% (2.3% to 14.3%)	0.009
Ventilator-free days at day 28	21.0 (0.0 - 25.0)	15.0 (0.0 - 22.2)	-6.0 (-8.1 to -3.9)	< 0.001
Duration of ventilation in survivors, days	6.0 (4.0 - 11.0)	8.0 (6.0 - 18.0)	2.0 (0.9 to 3.1)	< 0.001
Validation set				
ALVEOLI	n = 336	n = 157		
60-day mortality - no. (%)	59 (17.6)	68 (43.3)	25.8% (17.7% to 33.8%)	< 0.001
90-day mortality - no. (%)	60 (18.1)	70 (45.5)	27.3% (19.2% to 35.5%)	< 0.001
Ventilator-free days at day 28	21.0 (4.8 - 25.0)	2.0 (0.0 - 19.0)	-19.0 (-20.8 to -17.2)	< 0.001
Duration of ventilation in survivors, days	7.0 [4.0,13.0]	11.0 (6.0 - 22.2)	4.0 (2.1 to 5.9)	< 0.001
ARMA	n = 279	n = 100		
60-day mortality - no. (%)	69 (24.8)	42 (42.0)	17.2% (6.9% to 27.5%)	0.002
90-day mortality - no. (%)	70 (25.5)	42 (42.0)	16.5% (6.0% to 26.9%)	0.003
Ventilator-free days at day 28	17.0 (0.0 - 24.0)	2.0 (0.0 - 19.0)	-15.0 (-18.6 to -11.4)	< 0.001
Duration of ventilation in survivors, days	7.0 (4.0 - 13.8)	11.0 (5.0 -18.0)	4.0 (1.5 to 6.5)	0.018
SAILS	n = 319	n = 188		
60-day mortality - no. (%)	80 (25.1)	60 (31.9)	6.8% (-1.2% to 14.9%)	0.119
90-day mortality - no. (%)	81 (25.4)	63 (33.5)	8.1% (0.0% to 16.3%)	0.063
Ventilator-free days at day 28	21.0 (0.0 - 25.0)	16.0 (0.0 - 23.0)	-5.0 (-7.3 to -2.7)	< 0.001
Duration of ventilation in survivors, days	6.0 (3.0 - 10.0)	8.0 (5.0 - 14.0)	2.0 (0.7 to 3.3)	< 0.001
ART	n = 211	n = 298		
28-day mortality - no. (%)	81 (38.4)	180 (60.4)	22.0% (13.4% to 30.7%)	< 0.001
Ventilator-free days at day 28	0.0 (0.0 - 17.0)	0.0 (0.0 - 7.8)	-0.0 (-1.0 to 1.0)	< 0.001
Duration of ventilation in survivors, days	12.0 (8.0 - 20.0)	13.5 (8.0 - 20.0)	2.0 (-0.3 to 4.2)	0.570
Data are median (quartile 25th - quartile 75th) or N (%). Difference is mean difference with (95% CI) for binomial variables and median difference with (95% CI) for continuous variables Abbreviations: CI is the confidence interval.				

Identification of Subphenotypes

After comparing the clinical characteristics of the clusters, each cluster was assigned to represent a distinct subphenotype of ARDS, with patients in cluster 1 assigned to

subphenotype A, and patients in cluster 2 assigned to subphenotype B. Using blood biomarker information available for a subset of patients from both ARMA and ALVEOLI, subphenotype B showed increased levels of pro-inflammatory markers when compared to subphenotype A (**Figure 2** and **eTables 8 and 9**).

DISCUSSION

This study successfully demonstrated that nine easily obtainable clinical variables: arterial pH, partial O2 pressure, creatinine, bilirubin, bicarbonate, mean arterial pressure, heart rate, respiratory rate, and FiO2 at the time of study enrollment can identify two distinct ARDS subphenotypes with different clinical and biologic characteristics as well as outcomes across the test and validation cohorts. There was good generalizability amongst diverse populations from multiple validation datasets with temporal and geographical differences.

It is understandable that researchers feel compelled to use as much information as possible to build robust models. This is supportable for two main reasons: (1) the well-known heterogeneity of complex syndromes such as ARDS and (2) the abundance of highly granular clinical data generated by electronic health records (EHRs). It is anticipated that analyzing this vast amount of data will provide new knowledge regarding disease mechanisms by enabling researchers to find plausible hidden patterns within the data [29]. However, this data-heavy approach has the potential drawback of using predictors which are not generally obtained in a time window prior to intervention, or worse yet, using variables that are not part of the routine standard of care for patients. The rationale of using fewer and easy to collect clinical variables is not new in the field of critical care. Prognostic models have already shown that it is indeed feasible to create meaningful models using fewer predictors [30,31].

Unfortunately, unlike supervised algorithms (e.g., regression analyses), unsupervised algorithms such as K-means clustering do not provide one straightforward and established metric to describe feature importance. In that sense, our approach of

testing multiple sets of variables was also meant to select features that were most likely to be relevant, serving as surrogate for the feature selection step normally employed in supervised algorithms. While each individual variable by itself may not be significantly different across sub-phenotypes, their interaction in the 9-dimensional space of our model may be relevant.

Our initial choices to define variables commonly found in the EHR at ARDS diagnosis was inspired by a recent report from the World Health Organization (WHO) which showed an enormous discrepancy of medical devices availability in a survey across 135 countries [29]. Recognizing this inconsistency is essential for widespread implementation of machine learning models regardless of varying availability of resources across countries and health systems [29]. The aim is to provide clinically relevant information within a defined and short period that might impact the delivery of effective interventions to the right patient population and to as many patients as possible [29].

Recently, Sinha *et al.* developed supervised-learning gradient boosted classifier models trained using 24 or 14 readily available clinical data elements to reproduce biomarker-derived subphenotypes which were previously identified by Calfee *et al.* [17]. Unlike Sinha *et al.*, who predicted previously identified subphenotypes, our study has identified two subphenotypes *de novo* using a small set of clinical variables.

Although the subphenotypes that we have identified and those that have been previously published look similar, our work is distinct from previous studies in several ways. We employed different training and validation datasets as well as a different and well-established unsupervised learning technique. Moreover, we utilized a process for selecting predictors which is not comparable to previous studies. Acknowledging these

differences is crucial. It would not be unexpected to assume that these deviations would be relevant enough to produce different subphenotypes [32]. However, the clinical, laboratory characteristics, and the clinical outcomes of our subphenotypes show that they are remarkably similar to subphenotypes found in previous papers, regardless of methodological differences.

At this point it is not possible to go beyond this comparative analysis, as there is no gold standard definition of ARDS subphenotypes [32]. Nonetheless, our work does provide robust evidence that ARDS does indeed have two subphenotypes that can be systematically identified, despite major differences in population assessed and methodological approach used compared with previous studies. It also reinforces that we should continue to explore the underlying biological pathways of such subphenotypes to find responders to new or previously tested therapies.

Our study has several strengths. First, it is the largest cohort of patients that has been studied to develop distinct subphenotypes of ARDS patients. Moreover, our validation cohort included patients from the ART trial, allowing us to validate our model in the contemporaneous population of a large international randomized clinical trial in addition to the ARDSnet studies used in other subphenotyping studies. Second, our subphenotyping model was developed exclusively on the training set and then validated across multiple separate datasets. Nevertheless, similar separation in mortality was seen between the two subphenotypes across all trials. Third, we used the K-means algorithm to identify our subphenotypes, and the results obtained with this technique can be easily interpreted by clinicians and implemented in clinical practice. Lastly, this is the first phenotyping study that has used easily available clinical variables to identify ARDS

phenotypes *de novo*, which allows for early identification of these patients in the clinical care at the bedside. Using this algorithm with a small number of routinely collected variables could enable our model to be applied in trials that either retrospectively or prospectively assess interventions targeted to each subphenotype.

This study also has limitations. First, we have developed our models exclusively on patients enrolled in clinical trials. Due to the strict inclusion and exclusion criteria of these clinical trials, the generalizability of these results needs to be evaluated in unselected ARDS populations. Although there are clear clinical and biomarker differences between the identified subphenotypes, the model's clinical utility needs to be prospectively validated and further investigated. Additionally, our biomarker analysis is limited to those patients in which the data was made publicly available by the study authors, but future collection of biomarker data in a prospective study will allow more robust understanding of the underlying biology and validation of the subphenotype model. Also, K-means clustering does not handle missing data, and no approach was used to impute missing values. However, the extremely low rate of missingness in our study makes this issue less relevant. Lastly, future work should analyze previous trials to identify possible differential treatment responses for the subphenotypes of ARDS patients identified in this study.

CONCLUSIONS

This study confirms the existence of two distinct subphenotypes in ARDS patients using a novel clustering model on routinely collected clinical data. This work may allow for easier identification of ARDS subphenotypes to facilitate implementation of precision clinical trial

enrollment and development of targeted therapies in a variety of settings without the added burdens of biomarker evaluation.

DECLARATIONS

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing Interest: AD, MS, FGZ, ABC, ISM, DMP, LNL declare no relevant financial conflicts of interest. RK, EVA, LB, JO, DR and ROD are employees of Endpoint Health, Inc. ASN reported receiving personal fees from Dräger unrelated to the submitted work.

Ethics Approval: All patients provided informed consent in the original trials. This secondary analysis study was exempt from IRB review because it does not meet the definition of human subject as defined in 45 CFR 46.102. WIRB Work Order #1-1228617-1

Availability of data and material: Data from the ARDSnet studies (EDEN, FACTT, ARMA, ALVEOLI) may be requested from the NIH-NHLBI BioLINCC repository. These are available by application through the BioLINCC website (https://biolincc.nhlbi.nih.gov/resource_overview/), subject to established review processes for their distribution to qualified investigators and data from the ART trial can be requested from study authors.

Author Contributions: AD, RK, EVA, LB participated in study design and analysis, drafted, and revised the manuscript, and are the guarantor of the document. MS, DR, JO, FGZ, ABC, ISM, DMP, LNL, and ASN participated in interpretation of data analysis, drafted the manuscript, and revised it for critically important intellectual content. ROD participated in the study design, analysis, interpretation of data analysis, and final revision of the manuscript content.

Twitter: @msiuba, @f_g_zampieri, @rod_deliberato, @a_serpaneto, @l_bulgarelli, @endpointhealth

REFERENCES

1 ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, *et al.* Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;**307**:2526–33. doi:10.1001/jama.2012.5669

2 Thille AW, Esteban A, Fernández-Segoviano P, *et al.* Comparison of the Berlin Definition for Acute Respiratory Distress Syndrome with Autopsy. *Am J Respir Crit Care Med* 2013;**187**:761–7. doi:10.1164/rccm.201211-1981OC

3 Reilly J, Calfee C, Christie J. Acute Respiratory Distress Syndrome Phenotypes. *Semin Respir Crit Care Med* 2019;**40**:019–30. doi:10.1055/s-0039-1684049

4 Reddy K, Sinha P, O'Kane CM, *et al.* Subphenotypes in critical care: translation into clinical practice. *The Lancet Respiratory Medicine* 2020;**8**:631–43. doi:10.1016/S2213-2600(20)30124-7

5 Shankar-Hari M, Fan E, Ferguson ND. Acute respiratory distress syndrome (ARDS) phenotyping. *Intensive Care Med* 2019;**45**:516–9. doi:10.1007/s00134-018-5480-6

6 Calfee CS, Delucchi K, Parsons PE, *et al.* Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *The Lancet Respiratory Medicine* 2014;**2**:611–20. doi:10.1016/S2213-2600(14)70097-9

7 Famous KR, Delucchi K, Ware LB, *et al.* Acute Respiratory Distress Syndrome Subphenotypes Respond Differently to Randomized Fluid Management Strategy. *Am J Respir Crit Care Med* 2017;**195**:331–8. doi:10.1164/rccm.201603-0645OC

8 Calfee CS, Delucchi KL, Sinha P, *et al.* Acute respiratory distress syndrome subphenotypes and differential response to simvastatin: secondary analysis of a randomised controlled trial. *The Lancet Respiratory Medicine* 2018;**6**:691–8. doi:10.1016/S2213-2600(18)30177-2

9 for the NHLBI ARDS Network, Sinha P, Delucchi KL, *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**:1859–69. doi:10.1007/s00134-018-5378-3

10 Bos LD, Schouten LR, van Vught LA, *et al.* Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax* 2017;**72**:876–83. doi:10.1136/thoraxjnl-2016-209719

11 Ferguson ND, Fan E, Camporota L, *et al.* The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med* 2012;**38**:1573–82. doi:10.1007/s00134-012-2682-1

12 Bellani G, Laffey JG, Pham T, *et al.* Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA* 2016;**315**:788. doi:10.1001/jama.2016.0291

13 Gattinoni L, Vassalli F, Romitti F. Benefits and risks of the P/F approach. *Intensive Care Med* 2018;**44**:2245–7. doi:10.1007/s00134-018-5413-4

14 Matthay MA, Arabi YM, Siegel ER, *et al.* Phenotypes and personalized medicine in the acute respiratory distress syndrome. *Intensive Care Med* 2020;**46**:2136–52. doi:10.1007/s00134-020-06296-9

15 Shankar-Hari M, Rubenfeld GD. Population enrichment for critical care trials: phenotypes and differential outcomes. *Current Opinion in Critical Care* 2019;**25**:489–97. doi:10.1097/MCC.0000000000000641

16 Sinha P, Delucchi KL, McAuley DF, *et al.* Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *The Lancet Respiratory Medicine* 2020;**8**:247–57. doi:10.1016/S2213-2600(19)30369-8

17 Sinha P, Churpek MM, Calfee CS. Machine Learning Classifier Models Can Identify Acute Respiratory Distress Syndrome Phenotypes Using Readily Available Clinical Data. *Am J Respir Crit Care Med* 2020;**202**:996–1004. doi:10.1164/rccm.202002-0347OC

18 Kitsios GD, Yang L, Manatakis DV, *et al.* Host-Response Subphenotypes Offer Prognostic Enrichment in Patients With or at Risk for Acute Respiratory Distress Syndrome*: *Critical Care Medicine* 2019;**47**:1724–34. doi:10.1097/CCM.0000000000004018

19 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome. *N Engl J Med* 2000;**342**:1301–8. doi:10.1056/NEJM200005043421801

20 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Higher versus Lower Positive End-Expiratory Pressures in Patients with the Acute Respiratory Distress Syndrome. *N Engl J Med* 2004;**351**:327–36. doi:10.1056/NEJMoa032193

21 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network. Pulmonary-Artery versus Central Venous Catheter to Guide Treatment of Acute Lung Injury. *N Engl J Med* 2006;**354**:2213–24. doi:10.1056/NEJMoa061895

22 The National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Rice TW, Wheeler AP, *et al.* Initial Trophic vs Full

Enteral Feeding in Patients With Acute Lung Injury: The EDEN Randomized Trial. *JAMA: The Journal of the American Medical Association* 2012;**307**:795–803. doi:10.1001/jama.2012.137

23 The National Heart, Lung, and Blood Institute ARDS Clinical Trials Network. Rosuvastatin for Sepsis-Associated Acute Respiratory Distress Syndrome. *N Engl J Med* 2014;**370**:2191–200. doi:10.1056/NEJMoa1401520

24 Writing Group for the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial (ART) Investigators, Cavalcanti AB, Suzumura ÉA, *et al.* Effect of Lung Recruitment and Titrated Positive End-Expiratory Pressure (PEEP) vs Low PEEP on Mortality in Patients With Acute Respiratory Distress Syndrome: A Randomized Clinical Trial. *JAMA* 2017;**318**:1335. doi:10.1001/jama.2017.14171

25 Yehya N, Harhay MO, Curley MAQ, *et al.* Reappraisal of Ventilator-Free Days in Critical Care Research. *Am J Respir Crit Care Med* 2019;**200**:828–36. doi:10.1164/rccm.201810-2050CP

26 Castela Forte J, Perner A, van der Horst ICC. The use of clustering algorithms in critical care research to unravel patient heterogeneity. *Intensive Care Med* 2019;**45**:1025–8. doi:10.1007/s00134-019-05631-z

27 Ketchen DJ, Shook CL. The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. *Strategic Management Journal* 1996;**17**:441–58. doi:https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G

28 Caliński T, Harabasz J. A dendrite method for cluster analysis. *Communications in Statistics* 1974;**3**:1–27. doi:10.1080/03610927408827101

29 Bulgarelli L, Deliberato RO, Johnson AEW. Prediction on critically ill patients: The role of “big data.” *Journal of Critical Care* 2020;**60**:64–8. doi:10.1016/j.jcrc.2020.07.017

30 Johnson AEW, Kramer AA, Clifford GD. A New Severity of Illness Scale Using a Subset of Acute Physiology and Chronic Health Evaluation Data Elements Shows Comparable Predictive Accuracy*: *Critical Care Medicine* 2013;**41**:1711–8. doi:10.1097/CCM.0b013e31828a24fe

31 Deliberato RO, Escudero GG, Bulgarelli L, *et al.* SEVERITAS: An externally validated mortality prediction for critically ill patients in low and middle-income countries. *International Journal of Medical Informatics* 2019;**131**:103959. doi:10.1016/j.ijmedinf.2019.103959

32 DeMerle KM, Angus DC, Baillie JK, *et al.* Sepsis Subclasses: A Framework for Development and Interpretation. *Crit Care Med* Published Online First: 15 February 2021. doi:10.1097/CCM.0000000000004842

FIGURES LEGENDS

Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

Square symbols represent the study with the highest mean z score for each phenotype; Circles represent the study with the lowest mean z score for each phenotype. The colored bands are exclusively to help visualize the opposite trends of the variables on the different clusters; Art.pH: arterial pH; Bicarb: bicarbonate; MAP: mean arterial pressure; Creat: creatinine; Resp.Rate: respiratory rate

Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

For better visualization and due to difference in scales, the values were log-normalized and z-scored. Subphenotypes A and B are shown separately to highlight their differences.

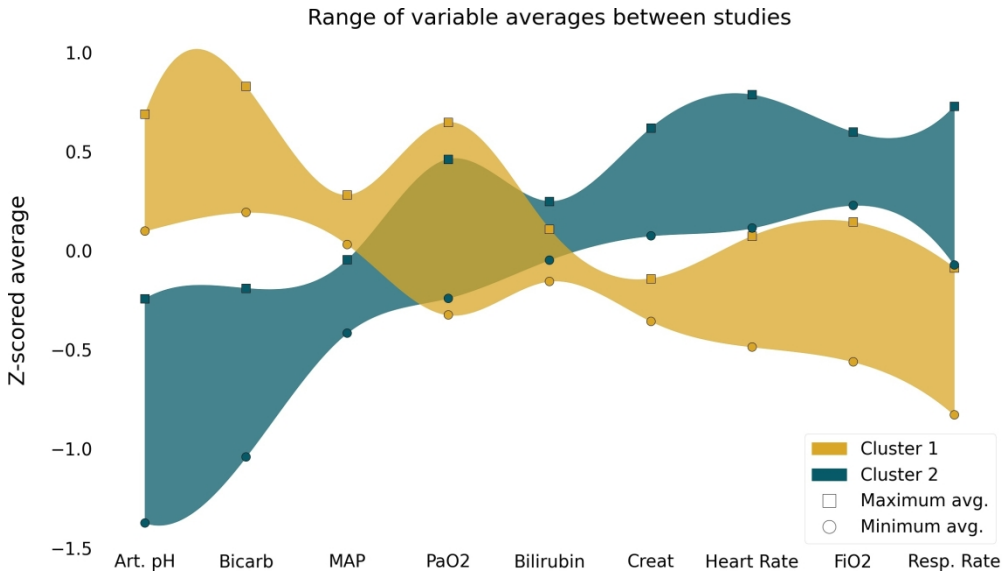


Figure 1 - Differences of the Variables Included in the Cluster Algorithm Among Clusters

812x457mm (72 x 72 DPI)

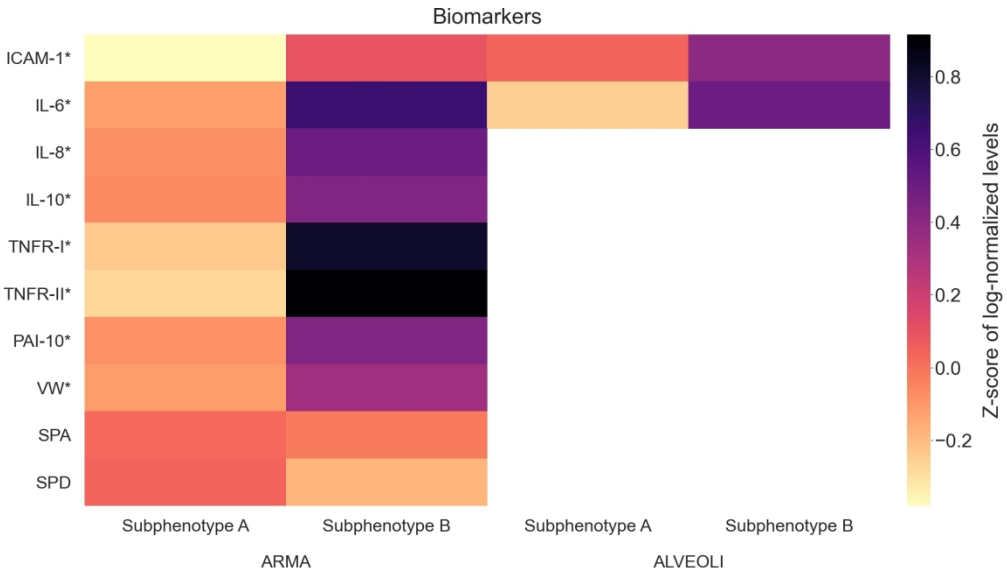


Figure 2 - Heat Map of the Biomarkers Available for the ARMA and ALVEOLI Trials

812x457mm (72 x 72 DPI)

Identification of Acute Respiratory Distress Syndrome subphenotypes de novo using routine clinical data: a retrospective analysis of ARDS clinical trials

ONLINE SUPPLEMENT

Additional Methods

Number of clusters

The optimal number of clusters was chosen according to two criteria: (1) Elbow method, by selecting a number of clusters that if further increased will result in only a small increase in performance and possibly cause overfit, hence this number is commonly referenced as to being in the "elbow" of the curve (**eFigure 1**); and (2) Calinski-Harabasz index, consisting of the ratio of *within* to *between* cluster dispersion; higher scores are indication of dense and well separated clusters (**e-Figure 1**).

Ventilator-free days

Ventilator free days for ALVEOLI, EDEN, FACTT, and SAILS were calculated according to the methods outlined by Yehya et al (1). Briefly, patients who died at any time in the 28 days were assigned 0 ventilator-free days. For survivors, the number of ventilator-free days was calculated based on the date of the final successful extubation; reintubations before the final extubation were not counted toward ventilator-free days. All days after a patient was discharged home up to the 28th day with unassisted breathing were assumed to be ventilator-free days.

eTable 1 - Percentage of missing data in the routinely collected variables, closest randomization, on EDEN and FACTT trials.

	EDEN (*n* = 1000)	FACTT (*n* = 999)
Age	0.0	0.0
Gender	0.0	0.0
Arterial pH	2.8	3.9
Bicarbonate	0.2	1.5
Bilirubin	8.1	26.8
Creatinine	0.0	0.0
FiO_2	0.8	0.6
Heart Rate	0.0	0.1
Height	0.1	0.9
Mean Arterial Pressure	12.1	0.8
$PaCO_2$	2.8	3.9
PaO_2	0.2	4.0
Positive end-expiratory pressure	1.0	0.3
Platelets	8.1	6.0
Plateau pressure	32.3	30.9
Respiratory rate	0.6	0.4
Tidal volume	15.3	12.1
Tidal volume per PBW	15.4	12.8

eTable 2 - Plausible physiological ranges for clinical measurements, closest to time of randomization

Variables	Lower Limit	Upper Limit
Age (years)	16	89
Arterial pH	6.65	7.80
Bicarbonate (mEq/L)	1	50
Bilirubin (mg/dL)	0.1	50
Creatinine (mg/dL)	0.1	20
FiO2	0.21	1
Heart Rate (beats per minute)	20	300
Height (cm)	120	220
Mean arterial pressure (mmHg)	10	400
PaCO2 (mmHg)	20	120
PaO2 / FiO2	0	500
PaO2 (mmHg)	30	500
PEEP (cm H20)	0	60
Platelets (thousands)	1	1000
Plateau Pressure (cm H20)	10	50
Respiratory Rate (resp per minute)	1	100
Tidal Volume (cm H20)	100	1400

eTable 3 - Correlation among fifteen routinely collected variables, close to the time of randomization.

	Age	pH	HCO_3	Bili	Creat	FiO_2	Gender	HR	MAP	$PaCO_2$	PaO_2	PEEP	Plat	RR	V_T/PBW
Age	1.00	0.06	-0.04	-0.02	0.11	-0.13	0.00	-0.27	-0.12	-0.11	-0.06	-0.22	0.00	-0.11	0.03
pH	0.06	1.00	0.40	-0.04	-0.16	-0.26	-0.01	-0.18	0.15	-0.39	0.00	-0.20	0.05	-0.21	0.07
HCO_3	-0.04	0.40	1.00	-0.08	-0.28	-0.05	-0.02	-0.18	0.08	0.44	0.02	-0.05	0.15	-0.24	-0.07
Bili	-0.02	-0.04	-0.08	1.00	0.06	-0.03	-0.04	0.01	-0.04	-0.01	0.03	0.01	-0.20	0.04	-0.01
Creat	0.11	-0.16	-0.28	0.06	1.00	-0.04	-0.08	-0.04	-0.01	-0.14	0.00	-0.06	-0.12	0.02	0.00
FiO_2	-0.13	-0.26	-0.05	-0.03	-0.04	1.00	0.03	0.13	-0.06	0.18	0.11	0.49	0.06	0.21	-0.02
Gender	0.00	-0.01	-0.02	-0.04	-0.08	0.03	1.00	-0.03	-0.05	-0.04	-0.06	0.02	0.09	0.09	0.19
HR	-0.27	-0.18	-0.18	0.01	-0.04	0.13	-0.03	1.00	-0.02	0.03	-0.04	0.12	-0.05	0.22	0.08
MAP	-0.12	0.15	0.08	-0.04	-0.01	-0.06	-0.05	-0.02	1.00	-0.03	0.01	-0.01	0.06	-0.04	0.00
$PaCO_2$	-0.11	-0.39	0.44	-0.01	-0.14	0.18	-0.04	0.03	-0.03	1.00	-0.04	0.17	0.11	-0.05	-0.17
PaO_2	-0.06	0.00	0.02	0.03	0.00	0.11	-0.06	-0.04	0.01	-0.04	1.00	-0.09	-0.04	-0.09	0.03
PEEP	-0.22	-0.20	-0.05	0.01	-0.06	0.49	0.02	0.12	-0.01	0.17	-0.09	1.00	0.00	0.33	-0.15
Plat	0.00	0.05	0.15	-0.20	-0.12	0.06	0.09	-0.05	0.06	0.11	-0.04	0.00	1.00	-0.05	0.03
RR	-0.11	-0.21	-0.24	0.04	0.02	0.21	0.09	0.22	-0.04	-0.05	-0.09	0.33	-0.05	1.00	-0.31
V_T/PBW	0.03	0.07	-0.07	-0.01	0.00	-0.02	0.19	0.08	0.00	-0.17	0.03	-0.15	0.03	-0.31	1.00

Data are Pearson correlation coefficients.

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 4 - Absolute 60-day Mortality Difference Among Clusters per Trial and Model

FACTT trial (*n* = 998)			EDEN trial (*n* = 1000)		
Model	**Patients scored***	**Mortality difference among clusters**	**Model**	**Patients scored***	**Mortality difference among clusters**
6	93.5%	13.1%	7	77.7%	8.1%
2	57.4%	12.5%	8	77.7%	8.1%
5	65.5%	12.2%	6	84.1%	6.7%
8	70.2%	11.6%	5	71.7%	6.5%
7	70.2%	11.4%	9	84.7%	6.1%
1	57.4%	11.2%	3	77.7%	4.4%
4	70.2%	10.6%	4	77.7%	4.0%
9	93.5%	10.4%	2	57.7%	3.9%
3	70.2%	10.1%	10	87.3%	2.8%
10	98.8%	3.9%	1	57.7%	0.1%

* Number of patients without any missing data, allowing their assignment to one of the clusters.

eTable 5 - List of variables in each model assessed

Model	Demographics		Arterial Blood Gases			Laboratory Values				Vital Signs			Ventilator Variables		
	Age	Gender	pH	PaO_2	$PaCO_2$	Creat	Bili	HCO_3	Plat	MAP	RR	HR	FiO2	PEEP	V_T/PBW
1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2			X	X	X	X	X	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X		X	X	X	X		
4	X	X	X	X		X	X	X		X	X	X	X		
5			X	X	X	X	X	X	X	X	X	X	X		
6	X	X	X	X		X		X		X	X	X	X		
7			X	X	X	X	X	X		X	X	X	X		
8			X	X		X	X	X		X	X	X	X		
9			X	X	X			X		X	X	X			
10	X	X								X	X	X			

Abbreviations: Bili denotes bilirubin, Creat is creatinine, HR is heart rate, MAP is mean arterial pressure, PEEP is positive end-expiratory pressure, Plat is platelets, RR is respiratory rate and V_T/PBW is tidal volume per predicted body weight.

eTable 6 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	ALVEOLI			ARMA		
	Cluster 1 (*n* = 336)	**Cluster 2 (*n* = 157)**	***p* value**	**Cluster 1 (*n* = 279)**	**Cluster 2 (*n* = 100)**	***p* value**
Age, year*	53.0 (39.0 - 66.2)	46.0 (37.0 - 60.0)	0.007	49.0 (37.0 - 64.0)	47.5 (36.0 - 61.0)	0.180
Male gender - no. (%)	188 (56.0)	86 (54.8)	0.883	169 (60.6)	61 (61.0)	0.965
Body mass index, kg/m^2	27.0 (22.9 - 31.1)	25.2 (21.7 - 30.2)	0.050	25.8 (23.0 - 30.2)	24.4 (21.5 - 29.7)	0.057
Caucasian - no. (%)	263 (78.3)	102 (65.0)	0.002	220 (78.9)	65 (65.0)	0.009
Etiology - no. (%)			0.001			< 0.001
Pneumonia	130 (38.7)	66 (42.0)		83 (29.7)	30 (30.0)	
Sepsis	63 (18.8)	50 (31.8)		64 (22.9)	43 (43.0)	
Aspiration	55 (16.4)	19 (12.1)		44 (15.8)	14 (14.0)	
Trauma	33 (9.8)	5 (3.2)		43 (15.4)	4 (4.0)	
Other	55 (16.4)	17 (10.8)		45 (16.1)	9 (9.0)	
Prognostic scores						
APACHE III	71. (59.0 - 83.0)	93.0 (80.0 - 110.0)	< 0.001	77.0 (66.0 - 90.5)	97.0 (81.8 (110.0)	< 0.001
Use of vasopressor - no. (%)	65 (20.1)	80 (51.3)	< 0.001	77 (27.6)	52 (52.5)	< 0.001
Vital signs						
Temperature, ºC	37.6 (37.1 - 38.2)	37.7 (36.9 - 38.3)	0.778	37.6 (37.1 - 38.1)	37.6 (36.8 - 38.4)	0.803
Heart rate, bpm	97.5 (83.0 - 109)	111.0 (97.0 - 126)	< 0.001	101.0 (89.0 - 112.5)	118 (105.0 - 128.0)	< 0.001
Mean arterial Pressure, mmHg	77.3 (77.0 - 87.3)	73.3 (65.0 - 80.3)	< 0.001	78.0 (70.7 - 88.0)	70.5 (64.9 - 80.4)	< 0.001
SpO_2, %	96 (94 - 97)	95 (92 - 97)	0.005	95 (93 - 98)	95.5 (93 - 97)	0.799
Urine output in 24 hours, mL	2065 (1355 - 3255)	1433 (569 - 2189)	< 0.001	2100 (1375 - 3096)	1525 (816 - 2650)	0.001
Laboratory tests						
Hematocrit, %	31.0 (28.0 - 34.0)	31.0 (27.0 - 35.0)	0.617	30.0 (28.0 - 33.0)	31.0 (28.0 - 34.0)	0.299
White blood cell count, $10^9/L$	11.7 (8.1 - 15.3)	10.7 (6.4 - 15.8)	0.166	11.9 (7.7 - 16.7)	9.8 (5.4 - 16.7)	0.057
Platelets, $10^9/L$	173 (94 - 266)	141 (57 - 214)	0.001	139 (80 - 212)	125 (72 - 196)	0.260
Creatinine, mg/dL	0.9 (0.7 - 1.3)	1.5 (0.9 - 3.0)	< 0.001	1.0 (0.7 - 1.4)	1.8 (1.2 - 3.2)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.9 (0.4 - 1.8)	0.289	1.0 (0.6 - 2.1)	1.1 (0.7 - 2.7)	0.106
Arterial blood gas						
pH*	7.42 (7.38 - 7.45)	7.31 (7.24 - 7.36)	< 0.001	7.42 (7.38 - 7.47)	7.33 (7.28 - 7.37)	< 0.001
PaO_2, mmHg	78 (68 - 93)	74 (65 - 92)	0.082	75 (66 - 91)	81 (68 - 96)	0.106
PaO_2 / FiO_2	149 (109 - 192)	103 (74 - 136)	< 0.001	118 (83 - 160)	99 (68 - 137)	0.006

$PaCO_2$, mmHg	38 (34 - 43)	36 (31 - 42)	0.046	37 (31 - 41)	34 (28.8 - 39.2)	0.003
Bicarbonate, mmol/L	24 (21 - 27)	17 (13 - 20)	< 0.001	23 (20 - 26)	16 (13 - 19)	< 0.001
Ventilatory variables						
Tidal volume, mL	500 (437 - 600)	480 (400 - 572)	0.002	700 (600 - 750)	700 (550 - 700)	0.198
Per PBW, mL/kg PBW	8.0 (6.9 - 9.5)	7.4 (6.2 - 9.2)	0.006	10.1 (9.2 - 11.1)	10.6 (9.0 - 11.4)	0.383
Plateau pressure, cmH_2O	25.0 (21.0 - 30.0)	29.0 (24.0 - 33.0)	< 0.001	29.0 (24.0 - 34.0)	31.0 (27.0 - 36.0)	0.018
PEEP, cmH_2O	10 (5 - 10)	10 (8 - 14)	< 0.001	8 (5 - 10)	10 (5 - 12)	0.150
Respiratory rate, breaths/min	20 (15 - 25)	30 (24 - 35)	< 0.001	18 (14 - 21)	24 (18.8 - 28)	< 0.001
FiO_2	0.50 (0.44 - 0.65)	0.75 (0.60 - 1.00)	< 0.001	0.60 (0.50 - 0.70)	0.70 (0.59 - 0.96)	< 0.001

Data are mean ± standard deviation, median (quartile 25^{th} - quartile 75^{th}) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight.

eTable 7 - Baseline Characteristics and Clinical Outcomes According to the Clusters and Two Trials in the Validation Set

	SAILS			ART		
	Cluster 1 (*n* = 319)	**Cluster 2 (*n* = 188)**	***p* value**	**Cluster 1 (*n* = 211)**	**Cluster 2 (*n* = 298)**	***p* value**
Age, year*	57.0 (46.0 - 67.0)	53.5 (39.0 - 65.0)	0.035	54.0 (37.0 - 65.0)	50.0 (35.2 - 61.0)	0.075
Male gender - no. (%)	150 (47.0)	100 (53.2)	0.211	136 (64.5)	181 (60.7)	0.448
Body mass index, kg/m^2	28.5 (23.9 - 34.6)	29.8 (23.2 - 35.1)	0.903	28.8 (24.6 - 35.6)	28.4 (25.0 - 31.7)	0.367
Caucasian - no. (%)	250 (78.4)	140 (74.5)	0.369	---	---	---
Etiology - no. (%)			0.709			0.052
Pneumonia	228 (71.5)	127 (67.6)		113 (53.6)	171 (57.4)	
Sepsis	63 (19.7)	39 (20.7)		38 (18.0)	59 (19.8)	
Aspiration	19 (6.0)	15 (8.0)		13 (6.2)	16 (5.4)	
Trauma	3 (0.9)	1 (0.5)		10 (4.7)	2 (0.7)	
Other	6 (1.9)	6 (3.2)		37 (17.5)	50 (16.8)	
Prognostic scores				---	---	---
APACHE III	70.0 (56.0 - 84.0)	92.0 (75.0 - 105.8)	< 0.001			
SAPS III	---	---	---	62 (50 - 71)	66 (53 - 75)	0.010
Use of vasopressor - no. (%)	150 (47.8)	142 (78.5)	< 0.001	130 (61.6)	242 (81.2)	< 0.001
Vital signs						
Temperature, ºC	37.2 (36.7 - 37.8)	37.3 (36.7 - 38.0)	0.346	---	---	---
Heart rate, bpm	91.0 (80.5 - 103.0)	102.0 (88.8 - 117.0)	< 0.001	90.0 (73.0 - 103.0)	112.0 (97.2 - 126.0)	< 0.001
Mean arterial Pressure, mmHg	78.0 (69.5 - 88.0)	70.0 (63.0 - 78.)	< 0.001	80.0 (73.5 - 89.0)	75.0 (70.0 - 83.0)	< 0.001
SpO_2, %	96 (95 - 99)	96 (93 - 99)	0.270	---	---	---
Urine output in 24 hours, mL	1570 (852 - 2383)	920 (350 - 1665)	< 0.001	---	---	---
Laboratory tests						
Hematocrit, %	31 (27 - 35)	31 (28 - 37)	0.142	---	---	---
White blood cell count, 10^9/L	13.6 (8.5 - 18.1)	15.4 (9.8 - 23.3)	0.009	---	---	---
Platelets, 10^9/L	164 (96 - 238)	131 (80 - 223)	0.032	177 (120 - 292)	169 (90 - 256)	0.048
Creatinine, mg/dL	1.0 (0.7 - 1.5)	1.4 (0.9 - 2.6)	< 0.001	1.0 (0.7 - 1.5)	1.7 (1.0 - 2.8)	< 0.001
Bilirubin, mg/dL	0.8 (0.5 - 1.4)	0.8 (0.5 - 1.6)	0.630	0.6 (0.4 - 1.2)	0.9 (0.4 - 1.7)	0.002
Arterial blood gas						
pH*	7.39 (7.35 - 7.44)	7.31 (7.24 - 7.35)	< 0.001	7.4 (7.3 - 7.4)	7.2 (7.2 - 7.3)	< 0.001
PaO_2, mmHg	82 (68 - 101)	86 (72 - 111.2)	0.112	118 (82 - 158)	104 (78 - 152)	0.065

PaO_2 / FiO_2	139 (98 - 195)	107 (74 - 159)	< 0.001	118 (82 - 158)	104 (78 - 152)	0.065
$PaCO_2$, mmHg	38 (34 - 45)	38 (32 - 44)	0.423	46 (41 - 56)	53 (42 - 65)	< 0.001
Bicarbonate, mmol/L	23 (20 - 26)	17 (14 - 21)	< 0.001	25.2 (22.5 - 28.8)	20.6 (17.8 - 23.4)	< 0.001
Ventilatory variables						
Tidal volume, mL	420 (360 - 480)	400 (340 - 450)	0.016	360 (320 - 400)	350 (300 - 397.8)	0.008
Per PBW, mL/kg PBW	6.4 (6.0 - 7.3)	6.1 (5.9 - 7.0)	0.030	6.0 (5.3 - 6.1)	5.9 (5.1 - 6.1)	0.034
Plateau pressure, cmH_2O	22.0 (18.0 - 27.0)	25.0 (20.0 - 29.0)	0.003	24.0 (21.0 - 28.0)	27.0 (23.0 - 30.0)	< 0.001
PEEP, cmH_2O	8 (5 - 10)	10 (8 - 13)	0.001	10 (10 - 14)	12 (10 - 14)	< 0.001
Respiratory rate, breaths/min	23 (19 - 27)	30 (24 - 35)	< 0.001	24 (20 - 28)	30 (24 - 34)	< 0.001
FiO_2	0.50 (0.40 - 0.60)	0.70 (0.50 - 0.90)	< 0.001	0.70 (0.60 - 0.80)	0.80 (0.70 - 1.00)	< 0.001

Data are mean ± standard deviation, median (quartile 25th - quartile 75th) or N (%)

Abbreviations: APACHE denotes Acute Physiology and Chronic Health Evaluation, V_T/PBW denotes tidal volume per predicted body weight...

eTable 8 - Biomarker levels by study and cluster

	ARMA				ALVEOLI			
	Subphenotype A (*n* = 279)	**Subphenotype B (*n* = 100)**	**Median Difference (95% CI)**	***p* value**	**Subphenotype A (*n* = 336)**	**Subphenotype B (*n* = 157)**	**Median Difference (95% CI)**	***p* value**
ICAM-1	654.0 (399.0 - 959.4)	888.0 (550.0 - 1365.3)	234 (60.3 to 407.8)	0.002	847.9 (585.7 - 1227.1)	1070.4 (748.2 - 1588.8)	219.4 (90.4 to 348.4)	< 0.001
IL-6	214.0 (91.8 - 553.5)	966.0 (291.0 - 2200.0)	749.1 (589.9 to 908.2)	< 0.001	182.5 (85.5 - 435.2)	775.0 (148.0 - 2846.5)	592 (515.5 to 668.6)	< 0.001
PAI-1	65.3 (37.8 - 109.5)	101.7 (50.8 - 291.6)	41 (18.3 to 63.7)	0.001	Not assessed	Not assessed	---	---
IL-8	46.0 (2.0 - 91.0)	106.9 (43.8 - 281.4)	60.9 (35.6 to 86.2)	< 0.001	Not assessed	Not assessed	---	---
IL-10	16.0 (0.0 - 40.3)	47.9 (0.0 - 120.7)	31.9 (20.2 to 43.6)	< 0.001	Not assessed	Not assessed	---	---
TNFR-I	2604.0 (1950.0 - 3777.0)	6897.0 (3622.5 - 12281.5)	4293 (3323.6 to 5262.4)	< 0.001	Not assessed	Not assessed	---	---
TNFR-II	6581.0 (4958.0 - 9658.0)	18611.0 (12262.5 - 35652.0)	12030 (9577.5 to 14482.5)	< 0.001	Not assessed	Not assessed	---	---
SPA	29.0 (11.8 - 68.0)	25.0 (10.5 - 40.0)	-4 (-19.9 to 11.9)	0.398	Not assessed	Not assessed	---	---
SPD	76.0 (36.2 - 145.2)	59.0 (30.0 - 125.0)	-18 (-52.6 to 16.6)	0.254	Not assessed	Not assessed	---	---
VW	308.0 (165.5 - 431.0)	384.0 (246.0 - 549.0)	76 (-26.5 to 178.5)	0.045	Not assessed	Not assessed	---	---

Data are median (quartile 25th - quartile 75th).

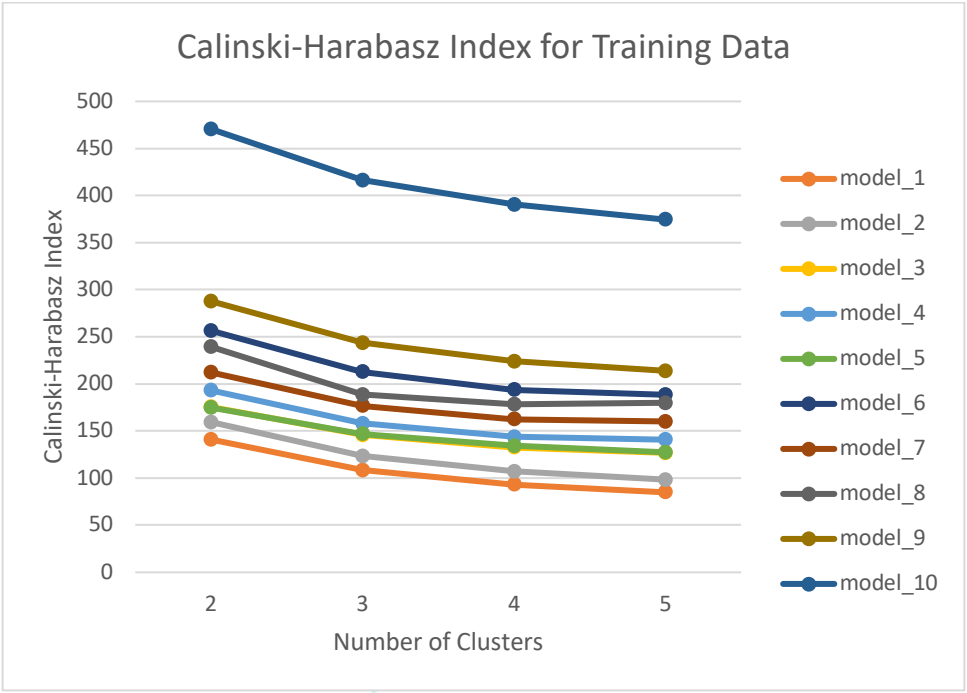
Abbreviations: 95%CI denotes 95% confidence interval, ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

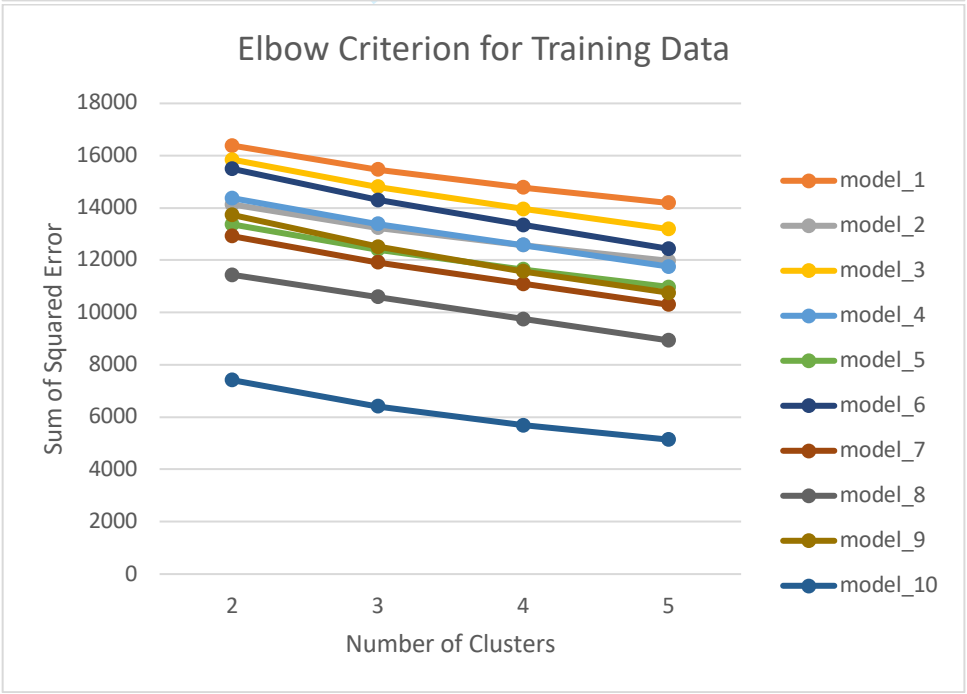
eTable 9 - Percentage of missingness in biomarker levels measured on day of randomization, on ARMA and ALVEOLI trials for patients with an assigned subphenotype

Biomarker	ARMA (*n* = 379)		ALVEOLI (*n* = 493)	
	Subphenotype A	Subphenotype B	Subphenotype A	Subphenotype B
ICAM-1	43%	31%	4%	3%
IL-6	41%	33%	4%	4%
PAI-1	42%	32%	Not assessed	Not assessed
IL-8	41%	33%	Not assessed	Not assessed
IL-10	42%	33%	Not assessed	Not assessed
TNFR-I	68%	61%	Not assessed	Not assessed
TNFR-II	68%	61%	Not assessed	Not assessed
SPA	67%	61%	Not assessed	Not assessed
SPD	67%	61%	Not assessed	Not assessed
VW	67%	61%	Not assessed	Not assessed

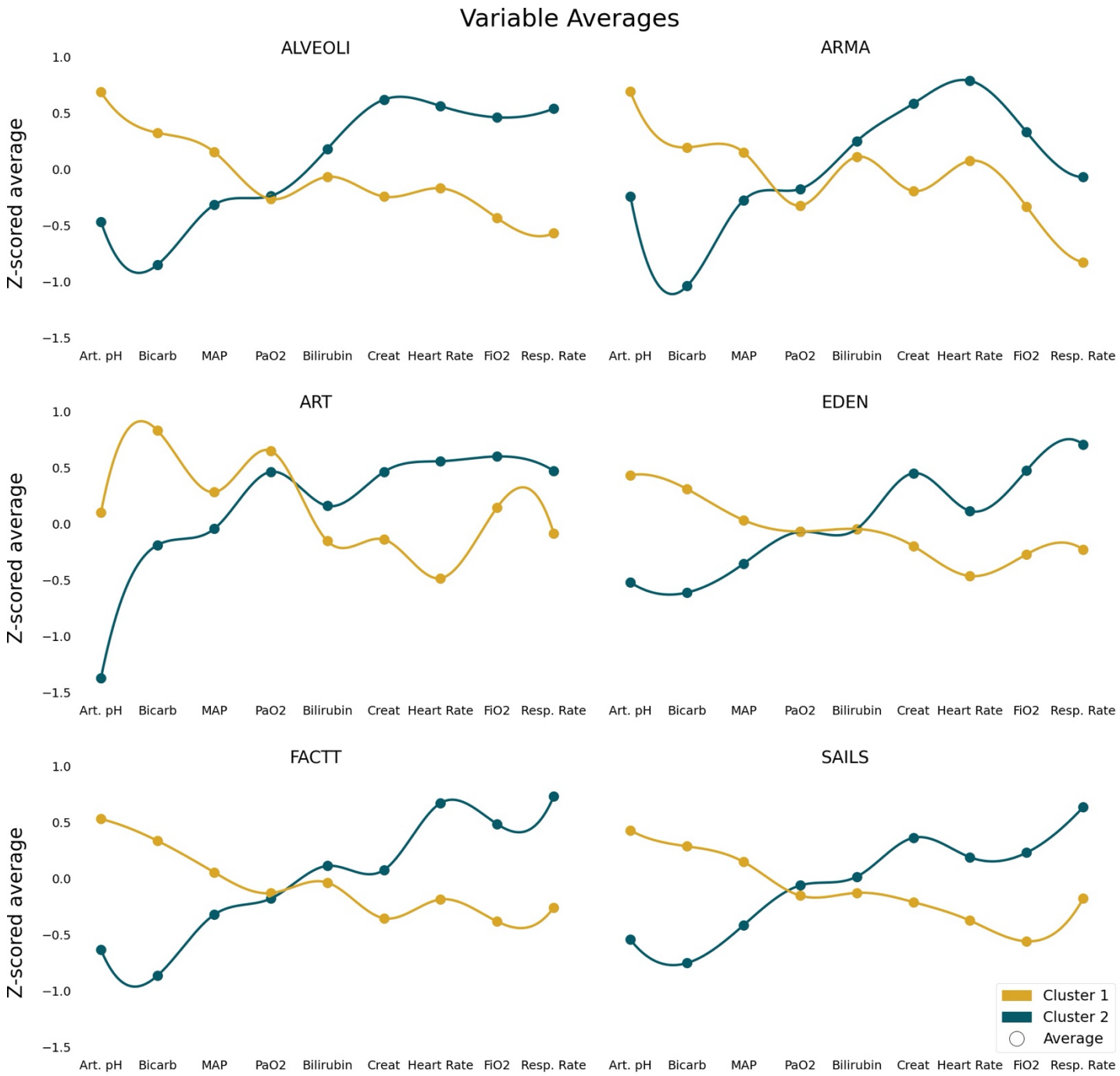
Abbreviations: ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is surfact protein A, SPD is surfact Protein D and VW is Von Willebrand factor.

eFigure 1 - Calinski-Harabasz Index and Elbow Method for Each of the 10 Models





eFigure 2 - Variable Averages for Each Study



The circles represent the averages for each variable. The colored lines are exclusively to help visualize the opposite trends of the variables on the different clusters.

Abbreviations: Art. pH is arterial pH, Bicarb is bicarbonate, MAP is mean arterial pressure, Creat is creatinine and Resp. Rate is respiratory rate

Reference:

1. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW. Reappraisal of Ventilator-Free Days in Critical Care Research. Am J Respir Crit Care Med.2019 Oct 1;200(7):828-836. doi: 10.1164/rccm.201810-2050CP.

Reporting checklist for prediction model development/validation.

Based on the TRIPOD guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the TRIPODreporting guidelines, and cite them as:

Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement.

		Reporting Item	Page Number
Title			
	#1	Identify the study as developing and / or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract			
	#2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction			
	#3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	6
	#3b	Specify the objectives, including whether the study describes the	6

		development or validation of the model or both.	
Methods			
Source of data	#4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8
Source of data	#4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	8
Participants	#5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8
Participants	#5b	Describe eligibility criteria for participants.	8
Participants	#5c	Give details of treatments received, if relevant	8
Outcome	#6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	9
Outcome	#6b	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	#7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured	8
Predictors	#7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	#8	Explain how the study size was arrived at.	8
Missing data	#9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	9
Statistical analysis methods	#10a	If you are developing a prediction model describe how predictors were handled in the analyses.	N/A
Statistical analysis methods	#10b	If you are developing a prediction model, specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	N/A
Statistical analysis methods	#10c	If you are validating a prediction model, describe how the predictions were calculated.	N/A
Statistical analysis	#10d	Specify all measures used to assess model performance and, if relevant,	10

methods		to compare multiple models.	
Statistical analysis methods	#10e	If you are validating a prediction model, describe any model updating (e.g., recalibration) arising from the validation, if done	N/A
Risk groups	#11	Provide details on how risk groups were created, if done.	11
Development vs. validation	#12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	10
Results			
Participants	#13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	12
Participants	#13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	12
Participants	#13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	12
Model development	#14a	If developing a model, specify the number of participants and outcome events in each analysis.	12
Model development	#14b	If developing a model, report the unadjusted association, if calculated between each candidate predictor and outcome.	N/A
Model specification	#15a	If developing a model, present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
Model specification	#15b	If developing a prediction model, explain how to the use it.	N/A
Model performance	#16	Report performance measures (with CIs) for the prediction model.	14
Model-updating	#17	If validating a model, report the results from any model updating, if done (i.e., model specification, model performance).	N/A
Discussion			
Limitations	#18	Discuss any limitations of the study (such as nonrepresentative sample,	19

		few events per predictor, missing data).	
Interpretation	#19a	For validation, discuss the results with reference to performance in the development data, and any other validation data	17
Interpretation	#19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	17
Implications	#20	Discuss the potential clinical use of the model and implications for future research	20
Other information			
Supplementary information	#21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	22
Funding	#22	Give the source of funding and the role of the funders for the present study.	22

This checklist was completed on 07. May 2021 using https://www.goodreports.org/, a tool made by the EQUATOR Network in collaboration with Penelope.ai