

**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 H2O)	0	60
Platelets (thousands)	1	1000
Plateau Pressure (cm H2O)	10	50
Respiratory Rate (resp per minute)	1	100
Tidal Volume (cm H2O)	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	FiO ₂	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, °C	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 ₂ O	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 ₂ O	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 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 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, °C	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 ₂ O	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 ₂ O	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).

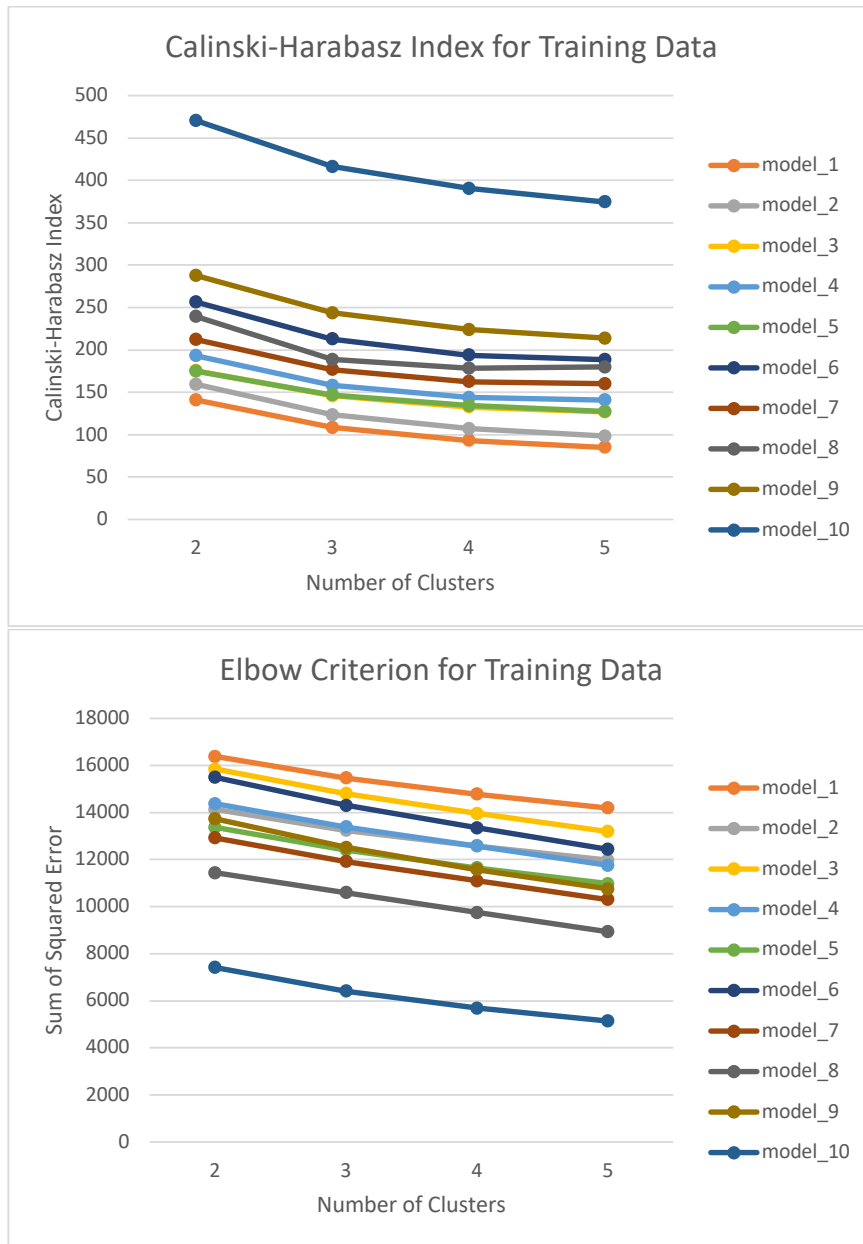
Abbreviations: 95%CI denotes 95% confidence interval, ICAM-1 is intercellular adhesion molecule-1, IL-6 is interleukin-6, PAI-1 is plasminogen activator inhibitor-1, IL-8 is interleukin-8, IL-10 is interleukin-10, TNFR-I is tumor necrosis factor receptor 1, TNFR-II is tumor necrosis factor II, SPA is 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

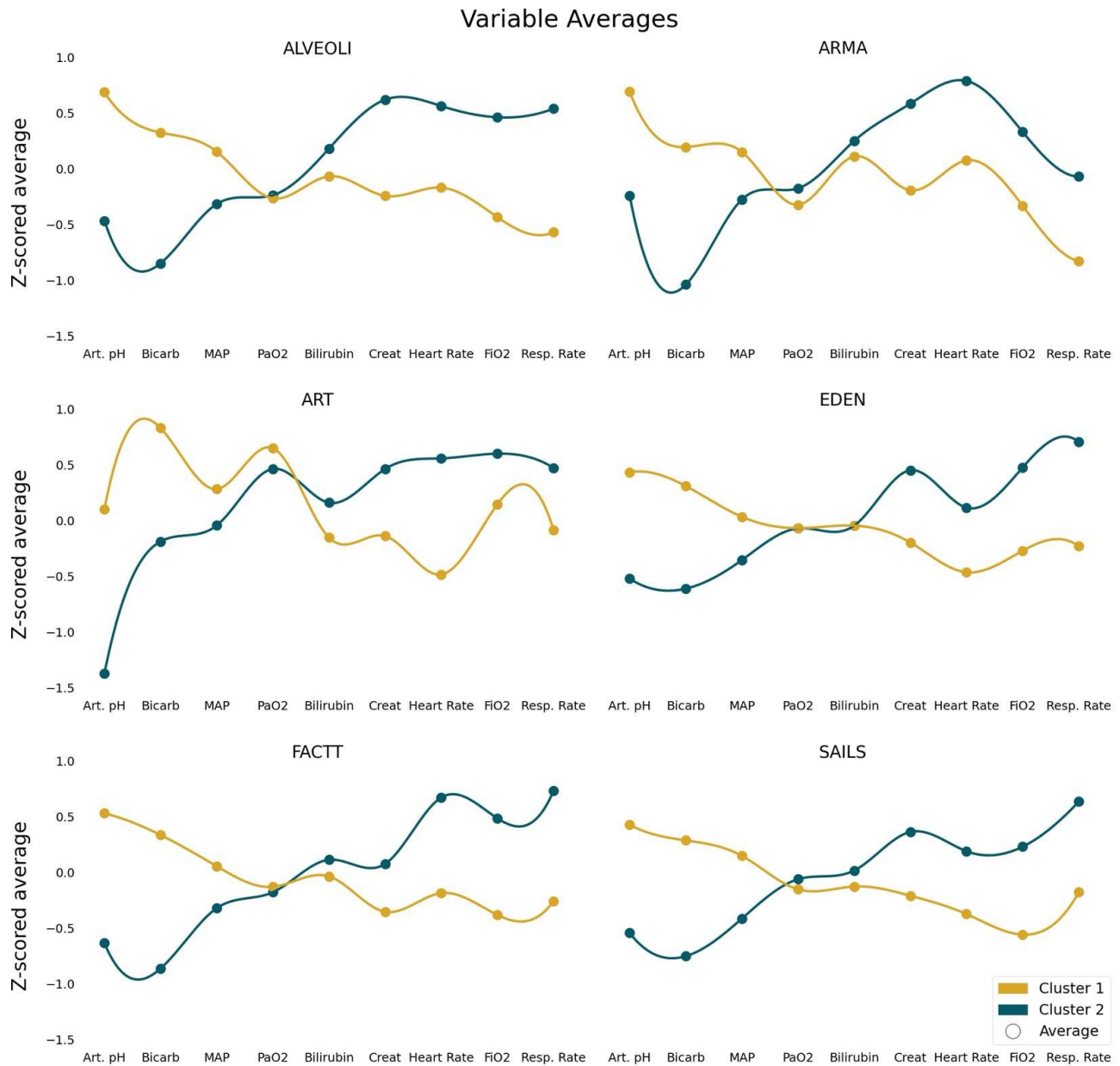
Biomarker	ARMA (n = 379)		ALVEOLI (n = 493)	
	Subphenotype A	Subphenotype B	Subphenotype A	Subphenotype B
ICAM-1	43%	31%	4%	3%
IL-6	41%	33%	4%	4%
PAI-1	42%	32%	Not assessed	Not assessed
IL-8	41%	33%	Not assessed	Not assessed
IL-10	42%	33%	Not assessed	Not assessed
TNFR-I	68%	61%	Not assessed	Not assessed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