Criteria Type	Criterion
General Inclusion	Age ≥ 18
	Mechanically ventilated in the ICU
General Exclusion	Transferred to ICU from outside facility
	Chronic tracheostomy prior to admission
	Known or suspected bronchopleural fistula
	Diffuse chronic fibrotic lung disease
	Pregnant women
	Prisoners
ARDS Patient Cohort	Dual clinician confirmed diagnosis of ARDS meeting
	Berlin criteria within 7 days of intubation
	Recorded lowest PaO2/FiO2 ratio ≤ 200 mm Hg
	Patients with chronic obstructive pulmonary disease and/or
	asthma were excluded from the ARDS patient cohort
Non-ARDS Patient Cohort	No suspicion of ARDS based on dual clinician adjudicated
	chart review

Supplemental Table 1. Inclusion/exclusion criteria for subject enrollment.

	ARDS	Non-ARDS
Primary Ventilator Mode (%)		
Assist Control-Pressure Control	60%	72%
Assist Control-Volume Control	26%	4%
Pressure-Regulated Volume Control	6%	2%
Pressure Support	6%	20%
Other	2%	2%
Mean RASS (range +4 to -5)	-3.1	-2.0
Cisatracurium Use (%)	26%	2%
Mean Duration of Cisatracurium (Hrs)	12.5	14.9

Supplemental Table 2. Additional clinical characteristics of the study cohort in the first 24 hours of mechanical ventilation.

Other, may include ventilator modes such as synchronized intermittent mandatory ventilation, volume support, and proportional assist ventilation; *RASS*, Richmond Agitation-Sedation Scale, range from +4 indicating combative to -5 indicating unarousable, with 0 indicating alert and calm

Feature Name	Units	Description
I-time	seconds	Total inspiratory time
E-time	seconds	Total expiratory time
I:E ratio	N/A	Ratio of I-time divided by E-time
Respiratory rate	breaths/min	Instantaneous respiratory rate, defined as 60/(I-time+E-time)
PEF to 0	NA	An expiratory time constant surrogate we defined by taking the slope of the expiratory flow from peak expiratory flow to where flow reaches close to 0.
PEF+0.16 to 0	NA	An expiratory time constant surrogate we defined by taking the slope of the expiratory flow from 0.16 seconds after peak expiratory flow to where flow reaches close to 0.
Mean expiratory flow	ml/min	The mean flow observation from the point in time peak expiratory flow (PEF) occurred to the point where the breath terminated and a new one began
Dynamic compliance (C_{dyn})	N/A	This measure is derived via: $C_{dyn} = \frac{TVi}{PIP-PEEP}$ where TVi is the inspiratory tidal volume. PIP is peak inspiratory pressure, and PEEP is positive end expiratory pressure.
Tidal volume ratio	N/A	Ratio of inspiratory tidal volume divided by expiratory tidal volume.

Supplemental Table 3. Features calculated for each breath in the analysis. Extraction code is publicly accessible at GitHub.

Algorithm	Sensitivity	Specificity	PPV	NPV	AUC
Random Forest	0.90±0.059	0.71±0.089	0.77±0.082	0.90±0.059	0.88±0.064
Neural Network	0.95±0.043	0.73±0.087	0.80±0.078	0.95±0.043	0.90±0.059
Adaboost	0.92±0.053	0.74±0.086	0.80±0.078	0.92±0.053	0.90±0.059
Logistic Regression	0.96±0.038	0.74±0.086	0.81±0.077	0.96±0.038	0.90±0.059
Naïve Bayes	0.82±0.075	0.80±0.078	0.82±0.075	0.85±0.070	0.89±0.061
SVM	0.98±0.027	0.68±0.091	0.77±0.082	0.98±0.270	0.89±0.061

Supplemental Table 4. Performance of different machine learning algorithms for ARDS classification. All numbers are rounded to 2 significant digits and reported along with 95% confidence intervals. *PPV*, positive predictive value; *NPV*, negative predictive value; *AUC*, area under the curve.

Split-Type	Model	N Trees	Max Tree Depth	Split Criterion
K-Fold	Train 24	33	2	Gini
Holdout	Train 24	5	6	Information Gain
Bootstrap	Train 24	33	2	Gini

Supplemental Table 5. Hyperparameters used for our Random Forest algorithm in accordance with our hyperparameter search methodology outlined above. We used the same hyperparameters for our bootstrap experiments as we did in k-fold experiments because the random nature of bootstrapping would have led to an un-fixed series of hyperparameters that changed each time we attempted to evaluate model hyperparameters. Other than the hyperparameters mentioned here, all other arguments were based on Scikit-learn Random Forest default arguments.

Feature	Average P-value	Average Rank
Mean flow from peak expiratory flow (PEF)	1.82e-45	1.0
Respiratory rate	2.25e-33	2.0
PEF to 0	9.88e-28	3.0
PEF+0.16 to 0	3.22e-22	4.0
I-time	2.21e-5	5.0
E-time	3.15e-4	6.0
I:E ratio	8.69e-3	7.0
Dynamic compliance (C_{dyn})	0.382	8.0
Tidal volume ratio	0.626	9.0

Supplemental Table 6. Chi^2-based feature importance based on train 24/test 24-hour (24/24) model dataset. Ranks are based on the average of all k-folds. Each feature achieved a whole number for a rank because there was no variation in feature ranking from fold to fold.

Feature	Average Score	Average Rank
Mean flow from peak expiratory flow (PEF)	0.355	1.0
Respiratory rate	0.2556	2.0
PEF to 0	0.1846	3.0
PEF+0.16 to 0	0.1038	4.0
I-time	0.0508	5.0
E-time	0.0298	6.0
I:E ratio	0.0127	7.0
Dynamic compliance (C_{dyn})	0.0061	8.0
Tidal volume ratio	0.0016	9.0

Supplemental Table 7. Feature importance using Gini importance values based on train 24/test 24-hour (24/24) model dataset. All scores were rounded to 4 significant digits. Ranks are based on the average of all k-folds. Each feature achieved a whole number for a rank because there was no variation in feature ranking from fold to fold. Note that feature rankings here were the same as they were in the chi^2 feature rankings.

Split-Type	Model	Features Used
K-Fold	Train 24 / Test 24 Hours	8
K-Fold	Train 24 / Test 6 Hours	3
Holdout	Train 24 / Test 24 Hours	5
Bootstrap	Train 24 / Test 24 Hours	7

Supplemental Table 8. Number of features selected for each model using the Random Forest algorithm. Feature selection was based on the AUC and accuracy selection method detailed in Supplemental Figure 3.

Model	Train/Test Split (n)	K-Fold Number	Sensitivity	Specificity	PPV	NPV	AUC
Train 24 / Test 6	80/14	1	1.0	0.59	0.74	1.0	0.89
-	-	2	0.96	0.85	0.88	0.95	0.97
-	-	3	0.71	0.67	0.74	0.65	0.7
-	-	4	0.82	1.0	1.0	0.83	0.99
-	-	5	1.0	0.66	0.77	1.0	0.88
-	NA	Mean of 5 k-folds	0.90±0.07	0.75±0.101	0.83±0.088	0.89±0.073	0.89±0.073

Supplemental Table 9. Performance statistics for the train 24/test 6-hour (24/6) model. Mean (with 95% confidence intervals) performance across all 5 k-folds is shown, and results of individual k-folds are displayed to illustrate the spectrum of performance variability. Note that only 70 subjects had VWD available in the 1st 6 hours resulting in a smaller sample size for the test cohort in the train 24/test 6-hour (24/6) model. *PPV*, positive predictive value; *NPV*, negative predictive value; *AUC*, area under the curve.

Split-Type	Sensitivity	Specificity	PPV	NPV	AUC
K-Fold	0.90±0.059	0.71±0.089	0.77±0.082	0.90±0.059	0.88±0.064
Holdout	0.90±0.107	0.75±0.155	0.79±0.146	0.89±0.112	0.94±0.085
Bootstrap	0.91±0.056	0.74±0.086	0.78±0.081	0.90±0.059	0.88±0.064

Supplemental Table 10. Comparative performance of k-fold, 70/30 holdout, and bootstrapping methods for our train 24/test 24-hour (24/24) model with the Random Forest algorithm. Results are displayed along with 95% confidence intervals. Note that confidence intervals for the 70/30 holdout split are wider than for k-fold and bootstrapping methods because only 30 subjects were used in the testing set, whereas bootstrapping and k-fold methods used all 100 subjects. *PPV*, positive predictive value; *NPV*, negative predictive value; *AUC*, area under the curve.