Distribution Agreement

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

Signature:

.

_____ _____

Joshua Detelich Date

Approval Sheet

Development of a Novel Risk Prediction Model in Acute Respiratory Distress Syndrome Utilizing Pulmonary Physiologic Parameters

By

Joshua Detelich
Master of Science
Clinical Research

_____
Greg S Martin, MD, MSc
Advisor

_____
Jordan Kempker, MD, MSc
Advisor

_____
Amita Manatunga, PhD
Committee Member

_____
Ashish Mehta, MD, MSc
Committee Member

Accepted:

_____
Kimberly Jacob Arriola, PhD, MPH
Dean of the James T. Laney School of Graduate Studies

_____
Date

Development of a Novel Risk Prediction Model in Acute Respiratory Distress Syndrome Utilizing Pulmonary Physiologic Parameters

By

Joshua Detelich
M.D., University of Texas Health Science Center at Houston, 2015
B.S., University of Texas at Austin, 2011

Advisors:
Greg S Martin, MD, MSc
Jordan A Kempker, MD, MSc

An abstract of
A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science in Clinical Research
2022

ABSTRACT

Development of a Novel Risk Prediction Model in Acute Respiratory Distress Syndrome Utilizing Pulmonary Physiologic Parameters
By Joshua Detelich

Introduction: Acute Respiratory Distress Syndrome (ARDS) is a condition that develops rapidly in response to a primary critical illness and leads to respiratory failure that requires mechanical ventilation to support. It has a mortality rate of 40% and there are no direct acting pharmacologic treatments. The limited interventions available are underutilized because clinicians lack a tool to predict early pulmonary outcomes for risk stratification and therapy selection. Our aim was to create a model predicting pulmonary worsening at 48 hours with a composite outcome of death or lack of improvement in both positive end expiratory pressure and fraction of inspired oxygen in a cohort of ARDS network trial participants.
Methods: We conducted a secondary data analysis of nine randomized control trials from the ARDS network. Participants were excluded if they were in a study arm that is no longer standard of care, had incomplete data or co-enrolled in more than one of the studies. Participants were randomly divided into derivation (70%) and validation (30%) cohorts. Multivariable logistic regression with automatic backward selection on readily available clinical, demographic, and pulmonary parameters was used to derive an initial model which was then refined through various methods. The final model was assessed using the area under a receiver operating curve (AUC) in both the derivation and validation cohort.
Results: The derivation cohort had 762 participants while the validation had 334. 461 (60.5%) of the participants in the derivation cohort experienced the outcome with only 4.1% due to death in the first 48 hours. The final derived model included oxygen saturation index, driving pressure, acute hepatic failure, history of hematologic malignancy and history of chronic pulmonary disease as its covariates. Efforts to refine the model made no significant improvements. The AUC was 0.643 on the derivation cohort and 0.641 on the validation set. A probability cutpoint of 0.56 could be used for a sensitivity of 76.8% and specificity of 37.5%.
Conclusions: A predictive model was created for a novel pulmonary outcome which used only readily available clinical parameters. However, it only had modest predictability and did not meet clinically significant sensitivity and specificity thresholds.

Development of a Novel Risk Prediction Model in Acute Respiratory Distress Syndrome Utilizing Pulmonary Physiologic Parameters

By

Joshua Detelich
M.D., University of Texas Health Science Center at Houston, 2015
B.S., University of Texas at Austin, 2011

Advisors:
Greg S Martin, MD, MSc
Jordan A Kempker, MD, MSc

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science in Clinical Research
2022

ACKNOWLEDGMENTS

This thesis project was only made possible through the support and aid from many key individuals. I would like to thank Dr. David Guidot, our Division Director, who has supported and invested in the development in my early career as an academic physician with sincere enthusiasm. I cannot overstate the contribution from my primary mentor, Dr. Jordan Kempker, who has been integral to the work at every step along the way from the design all the way up to analysis and presentation of results. He has always been available to help without hesitation and never stopped believing in the validity of my project. I also would like to thank Dr. Greg Martin, who helped mold my fledgling ideas into a cohesive aim with clinically meaningful goals. I want to acknowledge my committee members, Dr. Amita Manatunga and Dr. Ashish Mehta, for taking the time to oversee and review the presentation and writing of my thesis. Finally, to all the faculty and staff from the Master of Science in Clinical Research Program at Emory's Laney Graduate School, I am forever indebted to the invaluable knowledge I have gained during this training and carry it forward throughout my career.

The work and results from this thesis were derived using data from the ARDS Network Trials which was made available through the NHLBI BioLINCC Database. The work was supported through training grants from the National Center for Advancing Translational Sciences and NHLBI of the NIH under reward numbers UL1TR002378 and T32HL116271 respectively. The content presented here is solely the responsibility of the authors and does not reflect the views of the NIH or the ARDS Network.

TABLE OF CONTENTS

Page

A. INTRODUCTION.....1

B. BACKGROUND.....3

C. METHODS.....7

D. RESULTS.....12

E. DISCUSSION/CONCLUSIONS.....14

F. REFERENCES.....18

G. TABLES.....24

Table 1. Current Risk Stratification Models in ARDS.....24

Table 2. Physiologic Parameters and their Calculation.....25

Table 3. Summary of Covariates used to build logistic model.....26

Table 4. Demographics of excluded datasets and Initial cohort 26

Table 5. Baseline Characteristics of Initial Cohort 27

Table 6. Baseline Pulmonary Characteristics of Initial Cohort.....29

Table 7. Baseline Characteristics of Derivation Cohort 30

Table 8. Baseline Pulmonary Characteristics of Derivation Cohort.....32

Table 9. How Participants qualified for primary outcome in the derivation set.....34

Table 10. Final Logistic Model Parameters.....34

Table 11. Iterations of Model Refinement.....35

Table 12. Classification Table of the Final Model.....35

H. FIGURES.....36

Figure 1. Flow Diagram of Participant Selection into Final Cohorts.....36

Figure 2. Calibration Plot of Final Logistic Model.....36

Figure 3. Box and Whisker Plot of Final Logistic Model.....37

Figure 4. Receiver Operating Curves of Final Model.....38

INTRODUCTION

Acute Respiratory Distress Syndrome (ARDS) is a rapidly developing condition that leads to respiratory failure for which a patient requires a mechanical ventilator to support the functions of the pulmonary system (1). ARDS accounts for roughly 10% of all intensive care unit (ICU) patients and still today, has a mortality rate of 30-40% (2-4). Those that survive have significantly longer ICU stays associated with more severe morbidity compared to ICU patients that do not develop ARDS (3). Currently, there are no direct acting pharmacologic therapies, and all interventions are aimed at supporting pulmonary function while attempting to reduce rates of ventilator-induced injuries (1). Despite limited therapeutic options, there exists remarkable variation in practice and underutilization of even potentially lifesaving interventions such as prone ventilation (5). One potential explanation is that the clinical criteria to identify ARDS are sensitive but not specific, and therefore capture a patient population that includes those without ARDS who recover pulmonary function rapidly. The lack of any tool for a clinician to predict the early course for an ARDS patient may lead to a "wait and see" approach which delays implementation of any intervention to avoid exposing the patient to unnecessary risks and cost.

Current risk stratification models in ARDS all aim to predict mortality, and while some do perform quite well in this regard, they have not gained traction in stratifying clinical care early in a patient's trajectory (Table 1) (6-10). This could be from the observation that most patients with ARDS do not die directly from pulmonary failure but rather multiorgan failure after a prolonged ICU stay (11). Therefore, they do not help the clinician decide if more aggressive pulmonary interventions are warranted. A predictive

model focused on early pulmonary specific outcomes could help make this decision and overall improve utilization of these interventions.

The overall aim of this thesis project is to develop a model in a cohort of patients from the ARDS network trials that can predict those who will fail to have improvement in pulmonary status at 48 hours post diagnosis through a composite of death and change in positive end-expiratory pressure (PEEP) and fraction of inspired oxygen (FiO2) using readily available clinical variables including pulmonary physiologic parameters. The hypothesis for the final model is that early pulmonary worsening prediction would be aided by the inclusion of parameters of pulmonary physiology and a threshold of 80% sensitivity and 50% specificity could be reached. The model was created using multivariable logistic regression on a derivation cohort and tested against a separate validation cohort. The primary goal of the model is to aid clinicians in implementation of ARDS interventions and secondarily could potentially be used to aid future therapeutic trials research for cohort enrichment. Since the primary purpose was to aid clinicians, it was important that the final model be easily used clinically. As such, all variables included in the analysis are those readily available and previously described without any additional invasive testing required.

BACKGROUND

ARDS, first described in 1967, is a condition that only occurs secondary to a primary critical illness that acutely leads to pulmonary failure necessitating the use of mechanical ventilation for support (1). Generally, a patient will present for a life-threatening illness that causes a significant inflammatory response (i.e. sepsis, pneumonia, trauma, pancreatitis, etc.) and within a few days of onset, a maladaptive response to this inflammation damages the epithelial barrier of the alveoli, flooding the lung with proteinaceous fluid, profoundly altering respiratory mechanics and ultimately leading to inability to maintain adequate gas exchange (2). Nearly 25% of all ICU patients requiring mechanical ventilation develop or already have ARDS and still today, 30-40% of them die (3, 4).

Lung biopsy is the gold standard to diagnose ARDS, however since the risk of this is prohibitive in critically ill patients, clinicians rely on a set of clinical criteria which has evolved over the past 50 years. The current iteration is called the Berlin Criteria which was developed by an expert panel in 2012 (12, 13). The main goal of the panel was to improve clinical feasibility of the criteria to ultimately increase early recognition of ARDS. While the Berlin Criteria did make considerable efforts towards increasing identification of ARDS, it continued to rely on the ratio of dissolved oxygen in the blood to the fraction of inspired oxygen (PaO2/FiO2) as the sole predictor of severity by increased rates of mortality (12). Other risk stratification models have been created with improved performance, but still aim to predict overall mortality (Table 1) (6-10). Despite the improved predictability of these models, the Berlin Criteria remains the de facto risk stratification tool through use of PaO2/FiO2 for both clinicians and researchers alike due

to its simplicity.

Therapy for ARDS is limited and is focused on supporting the functions of the lung while reducing iatrogenic harm from the ventilator. Respiratory function is supported by the mechanical ventilator by offloading the workload from the patient and improving oxygen levels in the blood by increasing PEEP and FiO2. Ventilator induced lung injury can be limited through use of low tidal volume ventilation, a conservative fluid management strategy, neuromuscular blockade, and prone ventilation. There have been many direct pharmacologic agents tested, but to this day, none have shown clinically important outcomes (1). In a landmark study that shifted the paradigm of ARDS management, the ARMA trial showed that by using low tidal volume ventilation mortality was 31.0% compared to 39.8% (RR 0.78, p=0.007) in the control group (14). The FACTT trial showed that by limiting fluid administration, there was an increase in ventilator-free days (14.6±0.5 vs. 12.1±0.5, P<0.001) and ICU-free days (13.4±0.4 vs. 11.2±0.4, P<0.001) (15, 16). These two interventions carry no significant risk or increase in resource utilization, so have been widely accepted. In the PROSEVA trial, early implementation of prone ventilation demonstrated a reduction in mortality, with 16.0% vs 32.8% in the control group (HR 0.39; 95% CI 0.25-0.63; P<0.001; NNT=6) (17). Lastly, the ACURASYS trial found that with early neuromuscular blockade for 48 hours there was a mortality of 31.6% vs 40.7% in the control (RR 0.68; 95% CI 0.48-0.98; P=0.04; NNT=11) (18). These last two trials used a specific indication of PaO2/FiO2 < 150, but despite them both finding significant mortality improvement, there is overall underutilization of both interventions and more frequently they are implemented late as a rescue therapy rather than early in the clinical course (5). This can be partially explained by risks associated with each of these.

Prone ventilation positioning requires many trained staff members to turn the patients and avoid complications of pressure ulcers, physical injury, and accidental removal of vital catheters, chest tubes, and the endotracheal tube (19). Neuromuscular blockade significantly increases muscle weakness and requires the use of deep sedation with consequences of longer ventilation and more delirium (18). Considering these risks along with a lack of ability of clinicians to predict the early course of ARDS, it is understandable why there may be hesitation to implement these strategies as they were studied and instead opt for a "wait and see" approach. In fact, a recently published retrospective analysis found that over 10% of patients in some large, randomized control trials had resolved their ARDS within just one day of enrollment (20).

After being diagnosed with ARDS and placed on a mechanical ventilator, a patient usually follows one of three trajectories. Some will get better rapidly and continue a steady course of recovery as the aforementioned trial shows while others follow a opposite course of rapid decline towards death despite aggressive interventions. The majority lay somewhere in between however with perhaps and initial decline but then stabilize out and over the coming days to weeks either improve or not. By around 48-72 hours after diagnosis, which path a patient is on is usually clear. The patients with the in between pathway are the ones where theoretically there is a chance that early implementation of an intervention like prone ventilation may lead to their survival when they otherwise may have died. If there was a predictive model to find those that would have failure to improve in pulmonary status by 48 hours then these underutilized therapies could be started 1-2 days sooner than with the "wait and see" approach. The primary values that are evaluated multiple times throughout the day to see if pulmonary function is improving are the PEEP

and FiO2 settings on the ventilator, and if both are improving then it is safe to say that a patient has experienced pulmonary improvement.

The ideal model for this use case, would be simple and only utilize readily available clinical markers so that it can be easily adopted. Further, the model performance characteristics should favor sensitivity as to not miss a patient that may benefit from an intervention while tolerating a moderate false positive rate. Other than the usual clinical markers of any ICU patient, there have been many readily available pulmonary physiologic measures derived to measure the level of impairment in one of the three major categories of pulmonary physiology: oxygenation (21-26), dead-space ventilation (27-31), and compliance (32, 33). Oxygenation markers are surrogates for the amount of blood passing through the lungs inadequately oxygenated, dead-space ventilation measures aim to describe the amount of air going to areas in the lung with poor perfusion, and compliance markers are a way to understand the stiffness of the lungs. While all the pulmonary markers analyzed (Table 2) have been described previously in the literature, they have never been used in conjunction to create a predictive model.

METHODS

Hypothesis and Aim

Our aim for this study was to create a model predicting failure to have pulmonary improvement by 48 hours after enrollment into a cohort of randomized control trials from the ARDS network using only readily available clinical and pulmonary parameters. We hypothesized that such a model would aid clinicians in earlier and more equitable implementation of adjunctive therapies for ARDS such that it met sufficient sensitivity and specificity thresholds.

Study Design

We conducted secondary analyses of data from nine randomized controlled trials conducted by the ARDS Network from 1996-2013: ARMA, KARMA, LARMA, ALVEOLI, FACTT, ALTA, EDEN, OMEGA, and SAILS (14-16, 34-40). Data from these trials was obtained from the National Heart, Lung and Blood Institute (NHLBI) via the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC) program.

Study Characteristics

The ARDS network had an overarching goal of quickly and efficiently evaluating new therapeutics and interventions for patients with ARDS. As such, each trial shared many common inclusion and exclusion criteria. Included patients had to meet Berlin criteria for diagnosis of ARDS: 1) PaO2/FiO2 <300 while on mechanical ventilation, 2) bilateral infiltrates on chest imaging, 3) no clinical evidence of left atrial hypertension. Major exclusions were: 1) Age < 18 2) Neurologic conditions that would impair ability to wean from a ventilator, 3) severe chronic respiratory disease, 4) extensive burns, and 5)

organ transplant patients. Some trials allowed co-enrollment in a factorial design, and each is summarized below with more specific detail.

ARMA, KARMA, and LARMA were three prospective, multicenter, randomized control trials that utilized a factorial design to allow enrollment into ARMA plus either KARMA or LARMA. A total of 861 participants were enrolled between 1996-1999. ARMA compared a traditional ventilation strategy to one targeting lower tidal volume while KARMA compared ketoconazole to placebo and LARMA compared lysofylline to placebo. KARMA and LARMA found no significant effects in the treatment groups while ARMA found a mortality improvement in the low tidal volume group which set a new standard of care.

ALVEOLI was a prospective, multicenter, randomized control trial that enrolled 549 participants between 1999-2002 and compared a high PEEP strategy to a low PEEP strategy. This trial had a different age exclusion of < 13 years old and overall found no significant difference in the two ventilator strategies.

FACTT was a prospective, multicenter, randomized control trial that enrolled 1000 participants between 2000-2005 and compared a conservative vs liberal fluid management strategy. This trial had an age exclusion of < 13 years old and found that there was in increase in ventilator free days in those that received a conservative fluid management strategy which helped changed the standard of care.

ALTA was a prospective, multicenter, double-blind, randomized control trial that enrolled 282 participants between 2007-2008 and compared nebulized albuterol to placebo. This trial also allowed those 13 years of age and older. Overall, there was no significant difference found in the treatment groups.

EDEN and OMEGA were prospective, multicenter, randomized control trials with a factorial design that enrolled a total of 1000 participants from 2008-2011. EDEN compared early full-calorie enteral feeding to trophic enteral feeding while the OMEGA trial compared supplementation with n-3 fatty acids, y-linolenic acid, and antioxidants to placebo. There were no significant differences between groups in either trial.

SAILS was a prospective, multicenter, double-blind, randomized control trial that enrolled 745 participants from 2010-2013 and compared rosuvastatin to placebo. In addition to the common inclusion criteria, these participants had to have evidence of systemic inflammatory response syndrome with suspected infection. No significant treatment effect was found, and the trial stopped early due to futility.

Inclusions and Exclusions

All participants in the aforementioned trials were included in the initial cohort as long as they were not in any treatment arms that specifically contradict today's standard of care. From there, further exclusions removed datasets missing key variables used in the model creation, participants less than 18 years of age, and duplicates of participants co-enrolled in more than one study.

Independent Variables

The covariates used in the model derivation can be categorized into demographic information, clinical, and pulmonary parameters summarized in Table 3. Co-morbid conditions included were history of chronic lung disease, chronic dialysis, cirrhosis, solid tumor with metastasis, hematologic malignancy, diabetes, previous myocardial infarction, hypertension, congestive heart failure, prior stroke with sequelae, and immunosuppression. Baseline shock was defined as a mean arterial pressure less than 60 mmHg or use of a

vasopressor. Separate organ systems were defined as in failure by the following values: cardiac if systolic blood pressure <90 mmHg or on vasopressors, central nervous system if Glasgow coma score <12, coagulation if platelet < 80, renal if creatinine >2.0, and hepatic if bilirubin >2. Calculation of pulmonary parameters used is shown in greater detail in Table 2.

Outcomes

The primary outcome for our predictive model was pulmonary worsening at 48 hours. This was defined as a composite outcome of either death before 48 hours or lack of improvement in both PEEP and FiO2. Improvement in PEEP and FiO2 was defined as values at 48 hours being lower than the values at baseline or liberation from the ventilator before 48 hours. If either the PEEP or FiO2 was worse or stable, then that would meet criteria for the outcome.

A separate predictive model was also derived for a secondary outcome of 28-day mortality.

Analytic Plan

After applying inclusions and exclusions to the dataset, univariate analyses of the key descriptive variables were performed on this initial cohort. From here, any participants with outliers, impossible values, or missing independent variables were removed. The remaining cohort was split by simple random sampling with 70% going into a derivation cohort and 30% into a validation cohort. Bivariate analyses of the derivation cohort were performed with t-test, rank-sum, and chi-square tests appropriate. Multivariable logistic regression with automatic backward selection in SAS v9.4 on the derivation cohort was used to create the initial predictive model. A $p < 0.1$ was used as the selection criteria and

a complete case analytic strategy was used.

Model Refinement and Validation

The initial model underwent multiple steps of refinement as an attempt to improve its predictability. If there were multiple variables within the same pulmonary parameter category, then the one that resulted in the highest area under the receiver operator curve (AUC) was kept in order to avoid high collinearity. All continuous variables were transformed and evaluated in their square, cubic, and natural log forms to challenge the linearity assumption. A priori interaction terms of pulmonary parameters with body mass index (BMI) and history of chronic pulmonary disease were tested. A priori key variables of age, gender, BMI, and primary illness (cause of ARDS) were then forced back into the model if not already present. Finally, a LOESS smoother was used to find cut-points to transform any continuous variable into categorical. Each iteration was evaluated by likelihood ratio when models were nested with the original and otherwise an improvement in AUC of 0.05 was considered significant.

The fit of the final model was evaluated with the Hosmer-Lemeshow test, a calibration plot, and visually with a box and whisker plot. Overall model predictability was assessed with the AUC and tested against the validation cohort. Finally, a classification table of the final model was created to find an optimal probability cutpoint for the model use case.

RESULTS

Cohort Selection and Baseline Characteristics

There was a total of 5179 participants in the nine ARDS network trials available for analysis. All trials that occurred before ALTA did not have any variables available in the datasets for baseline co-morbid conditions and this was responsible for the majority of exclusions. A comparison of the basic demographics in this excluded cohort vs the analyzed is shown in Table 4. After further exclusions and simple random splitting of the cohort, there were 762 in the derivation and 334 in the validation cohorts (Figure 1). Baseline characteristics of the cohort before splitting and of the derivation cohort are shown in Tables 5-8. The cohort was relatively young with an average age (SD) of 53 (16), predominantly white (77%), and mainly living independently (81%). In terms of clinical variables, 44% were obese, 84% were medical ICU patients, 43% had baseline shock, 72% had moderate-severe ARDS by the Berlin definition, and pneumonia was the most common underlying condition at 64%. 461 out of the 762 in the derivation cohort had the primary outcome of pulmonary worsening at 48 hours and 35% experienced the secondary outcome of 28-day mortality. How those participants qualified for the composite primary outcome is shown in Table 9. Only 4.1% qualified by early mortality.

Primary Outcome Logistic Model Results

After applying automatic backward selection to the multivariable logistic regression model, the remaining covariates were oxygen saturation index, driving pressure, hepatic failure, presences of hematologic malignancy, and history of chronic pulmonary disease. The parameter coefficients are shown in Table 10. Efforts to refine the model made no significant difference in its predictability and all iterations are summarized in Table 11.

Validation of the model was first done by evaluating its fit to the data. A Hosmer-Lemeshow test was performed which showed a p-value of 0.53 indicating a failure to reject the null that the model was a good fit for the data. Fit was then evaluated visually with a calibration plot and box and whisker plot shown in Figures 2 and 3. The calibration plot also showed good fit and while the box and whisker plot had separation, there was significant overlap between those that did and did not get the outcome. A classification table was made and an optimal probability level of 0.56 led to accuracy of 61.3%, sensitivity of 76.8%, and specificity of 37.5% (Table 12). Finally, receiver operating curves for the derivation and validation cohort were overlayed in Figure 4 and resulted in AUC of 0.643 and 0.641 respectively.

Secondary Outcome Logistic Model Results

Following the same methodology used to arrive at the final model for the primary outcome a multivariate logistic regression model was derived for the secondary outcome of 28-day mortality. Its final covariates were oxygen saturation index, driving pressure, alveolar/arterial ratio, ventilatory ratio, age, gender, hepatic failure, baseline shock, coagulopathy, immunosuppression, and presence of hematologic malignancy. The AUC on the derivation and validation cohorts was 0.795 and 0.770 respectively. This was compared to the more general severity of illness calculator Acute Physiology, Age, and Chronic Health III (APACHE) (41) which was found to have an AUC of 0.685 and 0.683 on the derivation and validation cohorts respectively.

DISCUSSION

This study represents the first attempt at creation of a model in ARDS to predict an early pulmonary outcome. The final model predicting the primary outcome and consisting of oxygen saturation index, driving pressure, hepatic failure, presence of hematologic malignancy and history of chronic pulmonary disease had only modest predictability with an AUC of 0.643. The most optimal cutpoint for clinical use only yielded a sensitivity of 76.8% and specificity of 37.5%. The model did however fit the data well and performed nearly identically on the validation cohort. In terms of the secondary outcome of 28-day mortality, its final model consisted of the covariates: oxygen saturation index, driving pressure, alveolar/arterial ratio, ventilatory ratio, age, gender, presence of hematologic malignancy, immunosuppressed state, baseline shock, hepatic failure, and coagulopathy. It had moderate performance with AUC of 0.795 on the derivation cohort and 0.770 on the validation set.

Despite the ultimate result of our findings, the study had some key strengths. The cohort utilized allowed for a large sample size to be evaluated that all had rigorously collected randomized control trial data. In addition, there was a high enough event rate of the primary outcome to allow analysis of many covariates. Not only was this the first study to evaluate this novel pulmonary outcome, but it is also the first to combine this many previously described pulmonary physiologic parameters into a model building approach for ARDS. Finally, the derived model is simple enough to be used clinically. There are only 5 variables required with little to no chance of not having the information available on a routine ARDS patient. The only potential missing points would be an unknown history of hematologic malignancy or chronic pulmonary disease, and in these instances, it would

make reasonable clinical sense to assume the patient does not have these conditions.

Before fully interpreting the results of this study, it is critical to understand the limitations faced beyond the already expected limits of a secondary data analysis. A predictive model is only useful if it is able to perform well on a new cohort of patients, and there are two issues that make this model less generalizable. One is that all of the data used to derive the model is 1-2 decades old. ARDS management has not changed appreciably in that time however, so it is less likely to be a concern. Next, all of these participants were enrolled in rigorously designed controlled trials which has exclusions and rarely mimics real-life scenarios fully. The benefit of having reliable and consistently collected data to make the model initially though is likely better than having made a model from observational data. There were also many patients excluded from the original cohort and a high missingness in the data of a few covariates analyzed including driving pressure which made it into the final model. Despite this, there were still adequate patients and events available in the derivation and validation cohorts for the complete case analysis and when comparing the derivation set to the excluded population there were no significant differences in demographics (Table 4).

When interpreting the model, even though it did not perform well, the presence of these predictive covariates does make clinical sense. It is reasonable to believe that those with underlying lung disease would do worse with an acute lung ailment as is shown by the model. Also, as expected, pulmonary physiologic variables did play a role in predicting the outcome in all categories of pulmonary physiology except for dead space. This is unexpected because dead space estimates have repeatedly been shown to be associated with mortality in ARDS and is relevant in our 28-day mortality model (42). Hepatic failure and

history of hematologic malignancy are known to worsen outcomes in the ICU and are used in severity of illness scores such as the Simplified Acute Physiology Score II (SAPS II) and the APACHE III score (41, 43). Counterintuitively however, a higher oxygen saturation index and presence of hepatic failure are both predictive of lower odds of the outcome. This may be due to inherent bias in the outcome and could help explain why we did not develop a more clinically useful model which is explained in more detail below.

Our model aimed to predict a composite outcome of either death or failure to have improved PEEP and FiO2 by 48 hours of ARDS development. When interpreting the presence of oxygen saturation index in the model which has FiO2 in its calculation, it becomes clear that there is an inherent severity of illness bias within the outcome. Those participants that start at a higher baseline PEEP and FiO2 have more room for improvement than those with milder disease and starting already at a low value for each of these variables. This creates a scenario where the model is being created with the balance of two opposing forces: the true severity of illness by physiologic measures vs. the practical nature of the worse disease having more room for improvement. This finding means that even had the model performed better, its clinical utility would be questionable. Otherwise, this model illustrates that predictions in the critically ill patient are difficult. Even models derived on thousands of patients such as the APACHE scores perform well on some cohorts and significantly worse others as seen in our study (41, 44). We believed that by targeting an earlier outcome, fewer variations would come into consideration and therefore prediction should be easier, but that was not the case.

When evaluating the performance of the model for the secondary outcome, it had similar predictability to other models that have been created for mortality in patients with

ARDS (Table 1). In addition, while there were pulmonary physiologic parameters included, it was heavily driven by extrapulmonary markers such as age, baseline shock, and immunosuppression. This finding is expected given that most patients with ARDS do not die directly from pulmonary failure, but rather multiorgan dysfunction that develops over time from the underlying critical illness (11).

Conclusions and Future Directions:

This study developed a predictive model for a novel outcome of early pulmonary worsening in a cohort of ARDS patients. The combination of limited predictability and inherent severity of illness bias in the outcome measure make it unsuitable for clinical use in its current state. With the COVID-19 being a significant cause of ARDS, conscious and equitable resource utilization has been brought to the forefront. Now more than ever has a model such as the one aimed to be created in this work been more relevant and necessary for clinicians to have at their disposal. Further research should be done to create a model accurately predicting early pulmonary outcomes with readily available baseline clinical markers because of this strong need. It could be improved by changing the definition of pulmonary worsening, focusing on more severe patients, or creating separate models for different phenotypes of ARDS. Race was included as one of our evaluated covariates and ended up not being predictive but would have led to more discussion had it been. I believe future models should not include this as a covariate even if it is predictive due to the inherent inequity it may lead to. Finally, future methods should utilize capabilities of machine learning which is a growing field with significant promise if applied correctly. It also has the benefit of being able to continually improve as more data is fed into it which helps keep a model relevant to changing practices in medicine.

REFERENCES

1. Thompson BT, Chambers RC, Liu KD: Acute Respiratory Distress Syndrome. *N Engl J Med* 2017; 377(6):562-572

2. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, et al: Acute respiratory distress syndrome. *Nat Rev Dis Primers* 2019; 5(1):18

3. Kang M, Kempker JA: Definitions, Epidemiology, Clinical Risk Factors, and Health Disparities in Acute Respiratory Distress Syndrome. *Semin Respir Crit Care Med* 2019; 40(1):3-11

4. Bellani G, Laffey JG, Pham T, Fan E, et al: Epidemiology, Patterns of Care, and Mortality for Patients With Acute Respiratory Distress Syndrome in Intensive Care Units in 50 Countries. *JAMA* 2016; 315(8):788-800

5. Qadir N, Bartz RR, Cooter ML, Hough CL, et al: Variation in Early Management Practices in Moderate-to-Severe ARDS in the United States: The Severe ARDS: Generating Evidence Study. *Chest* 2021; 160(4):1304-1315

6. Santos RS, Silva PL, Rocco JR, Pelosi P, et al: A mortality score for acute respiratory distress syndrome: predicting the future without a crystal ball. *J Thorac Dis* 2016; 8(8):1872-1876

7. Kangelaris KN, Calfee CS, May AK, Zhuo H, et al: Is there still a role for the lung injury score in the era of the Berlin definition ARDS? *Ann Intensive Care* 2014; 4(1):4

8. Cooke CR, Shah CV, Gallop R, Bellamy S, et al: A simple clinical predictive index for objective estimates of mortality in acute lung injury. *Crit Care Med* 2009; 37(6):1913-1920

9. Villar J, Ambrós A, Mosteiro F, Martínez D, et al: A Prognostic Enrichment Strategy for Selection of Patients With Acute Respiratory Distress Syndrome in Clinical Trials. *Crit Care Med* 2019; 47(3):377-385

10. Villar J, Ambrós A, Soler JA, Martínez D, et al: Age, PaO2/FIO2, and Plateau Pressure Score: A Proposal for a Simple Outcome Score in Patients With the Acute Respiratory Distress Syndrome. *Crit Care Med* 2016; 44(7):1361-1369

11. Villar J, Martínez D, Mosteiro F, Ambrós A, et al: Is Overall Mortality the Right Composite Endpoint in Clinical Trials of Acute Respiratory Distress Syndrome? *Crit Care Med* 2018; 46(6):892-899

12. Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, et al: Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012; 307(23):2526-2533

13. Ferguson ND, Fan E, Camporota L, Antonelli M, et al: The Berlin definition of ARDS: an expanded rationale, justification, and supplementary material. *Intensive Care Med* 2012; 38(10):1573-1582

14. Brower RG, Matthay MA, Morris A, Schoenfeld D, et al: 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(18):1301-1308

15. Wheeler AP, Bernard GR, Thompson BT, Schoenfeld D, et al: Pulmonary-artery versus central venous catheter to guide treatment of acute lung injury. *N Engl J Med* 2006; 354(21):2213-2224

16. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, et al: Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med* 2006; 354(24):2564-2575

17. Guérin C, Reignier J, Richard JC, Beuret P, et al: Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 2013; 368(23):2159-2168

18. Papazian L, Forel JM, Gacouin A, Penot-Ragon C, et al: Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med* 2010; 363(12):1107-1116

19. Aoyama H, Uchida K, Aoyama K, Pechlivanoglou P, et al: Assessment of Therapeutic Interventions and Lung Protective Ventilation in Patients With Moderate to Severe Acute Respiratory Distress Syndrome: A Systematic Review and Network Meta-analysis. *JAMA Netw Open* 2019; 2(7):e198116

20. Schenck EJ, Oromendia C, Torres LK, Berlin DA, et al: Rapidly Improving ARDS in Therapeutic Randomized Controlled Trials. *Chest* 2019; 155(3):474-482

21. Balzer F, Menk M, Ziegler J, Pille C, et al: Predictors of survival in critically ill patients with acute respiratory distress syndrome (ARDS): an observational study. *BMC Anesthesiol* 2016; 16(1):108

22. Monchi M, Bellenfant F, Cariou A, Joly LM, et al: Early predictive factors of survival in the acute respiratory distress syndrome. A multivariate analysis. *Am J Respir Crit Care Med* 1998; 158(4):1076-1081

23. DesPrez K, McNeil JB, Wang C, Bastarache JA, et al: Oxygenation Saturation Index Predicts Clinical Outcomes in ARDS. *Chest* 2017; 152(6):1151-1158

24. Chen WL, Lin WT, Kung SC, Lai CC, et al: The Value of Oxygenation Saturation Index in Predicting the Outcomes of Patients with Acute Respiratory Distress Syndrome. *J Clin Med* 2018; 7(8)

25. Metkus TS, Tampakakis E, Mullin CJ, Houston BA, et al: Pulmonary Arterial Compliance in Acute Respiratory Distress Syndrome: Clinical Determinants and

Association With Outcome From the Fluid and Catheter Treatment Trial Cohort. *Crit Care Med* 2017; 45(3):422-429

26. Chen W, Ware LB: Prognostic factors in the acute respiratory distress syndrome. *Clin Transl Med* 2015; 4(1):65

27. Sinha P, Calfee CS, Beitler JR, Soni N, et al: Physiologic Analysis and Clinical Performance of the Ventilatory Ratio in Acute Respiratory Distress Syndrome. *Am J Respir Crit Care Med* 2019; 199(3):333-341

28. Morales-Quinteros L, Schultz MJ, Bringué J, Calfee CS, et al: Estimated dead space fraction and the ventilatory ratio are associated with mortality in early ARDS. *Ann Intensive Care* 2019; 9(1):128

29. Gajic O, Afessa B, Thompson BT, Frutos-Vivar F, et al: Prediction of death and prolonged mechanical ventilation in acute lung injury. *Crit Care* 2007; 11(3):R53

30. Siddiki H, Kojicic M, Li G, Yilmaz M, et al: Bedside quantification of dead-space fraction using routine clinical data in patients with acute lung injury: secondary analysis of two prospective trials. *Crit Care* 2010; 14(4):R141

31. Beitler JR, Thompson BT, Matthay MA, Talmor D, et al: Estimating dead-space fraction for secondary analyses of acute respiratory distress syndrome clinical trials. *Crit Care Med* 2015; 43(5):1026-1035

32. Amato MB, Meade MO, Slutsky AS, Brochard L, et al: Driving pressure and survival in the acute respiratory distress syndrome. *N Engl J Med* 2015; 372(8):747-755

33. Guerin C, Papazian L, Reignier J, Ayzac L, et al: Effect of driving pressure on mortality in ARDS patients during lung protective mechanical ventilation in two randomized controlled trials. *Crit Care* 2016; 20(1):384

34. Network TANAftA: Ketoconazole for early treatment of acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. The ARDS Network. *JAMA* 2000; 283(15):1995-2002

35. Network TACT: Randomized, placebo-controlled trial of lisofylline for early treatment of acute lung injury and acute respiratory distress syndrome. *Crit Care Med* 2002; 30(1):1-6

36. Brower RG, Lanken PN, MacIntyre N, Matthay MA, et al: Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Engl J Med* 2004; 351(4):327-336

37. Rice TW, Wheeler AP, Thompson BT, deBoisblanc BP, et al: Enteral omega-3 fatty acid, gamma-linolenic acid, and antioxidant supplementation in acute lung injury. *JAMA* 2011; 306(14):1574-1581

38. Rice TW, Wheeler AP, Thompson BT, Steingrub J, et al: Initial trophic vs full enteral feeding in patients with acute lung injury: the EDEN randomized trial. *JAMA* 2012; 307(8):795-803

39. Matthay MA, Brower RG, Carson S, Douglas IS, et al: Randomized, placebo-controlled clinical trial of an aerosolized β_2-agonist for treatment of acute lung injury. *Am J Respir Crit Care Med* 2011; 184(5):561-568

40. Truwit JD, Bernard GR, Steingrub J, Matthay MA, et al: Rosuvastatin for sepsis-associated acute respiratory distress syndrome. *N Engl J Med* 2014; 370(23):2191-2200

41. Knaus WA, Wagner DP, Draper EA, Zimmerman JE, et al: The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991; 100(6):1619-1636

42. Nuckton TJ, Alonso JA, Kallet RH, Daniel BM, et al: Pulmonary dead-space fraction as a risk factor for death in the acute respiratory distress syndrome. *N Engl J Med* 2002; 346(17):1281-1286

43. Le Gall JR, Lemeshow S, Saulnier F: A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993; 270(24):2957-2963

44. Pirracchio R, Petersen ML, Carone M, Rigon MR, et al: Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med* 2015; 3(1):42-52

TABLES

Table 1. Current Risk Stratification Models in ARDS

Scoring system	Parameters used	Performance
Berlin Criteria	PaO2/FiO2	Mortality VFD(IQR) Mild 30% 16(0-24) Moderate 35% 11(0-21) Severe 41% 0(0-18) Mortality AUC 0.58
Lung Injury Score	Chest X-ray, PaO2/FiO2, PEEP, Pulmonary Compliance	Mortality AUC 0.58
Villar et al.	Age, PaO2/FiO2, Plateau Pressure, Extrapulmonary organ failure	Mortality AUC 0.86
Cooke et al.	Hematocrit Serum Bilirubin Net Fluid Balance Age	Mortality AUC 0.72

Abbreviations: PaO2/FiO2: Ratio of blood oxygen to fraction of inspired oxygen. PEEP: positive end expiratory pressure. VFD: Ventilator free days. AUC: Area under curve

Table 2. Physiologic Parameters and their Calculation

Physiologic Category	Parameter	Formula
Oxygenation	Oxygenation Index	$\frac{MAPxFiO2x100}{PaO2}$
	Oxygen Saturation Index	$\frac{MAPxFiO2x100}{SpO2}$
	Alveolar-arterial Gradient	$PAO2 - PaO2$
	Arterial-Alveolar Ratio	$\frac{PaO2}{PAO2}$
Dead Space Ventilation	Ventilatory Ratio	$\frac{VExPaCO2}{VEpredx37.5}$
	Minute Ventilation for PaCO2=40	$\frac{VExPaCO2}{40}$
	Siddiki Estimate	$1 - \frac{\left(\frac{0.86x(HBxhfx0.8)}{6.8644}\right)}{VExPaCO2}$
	Harris-Benedict Estimate	$1 - \frac{0.86xHB/8.604}{VExPaCO2}$
Compliance	Pulmonary Compliance	$\frac{Vt}{(Pplat - PEEP)}$
	Driving Pressure	$Pplat - PEEP$

Abbreviations: MAP:Mean Airway pressure, FiO2: Fraction of inspired oxygen, SpO2: Oxygen saturation of hemoglobin, PAO2: partial pressure of oxygen in alveoli, PaO2: partial pressure of oxygen in blood, VE: Minute ventilation, PaCO2: partial pressure of carbon dioxide in blood, VEpred: Predicted minute ventilation, HB: Harris Benedict basal metabolic rate, hf: hypermetabolic correction factor, Vt: Tidal Volume, Pplat: Plateau pressure, PEEP: Positive end expiratory pressure

Table 3. Summary of Covariates used to build logistic model

Category	Variables
Demographics	Age Gender Race BMI Place of residence
Clinical	Co-morbidities Shock Vasopressors Organ Failure Primary Illness
Pulmonary Parameters	Oxygenation index Saturation index A-a ratio and gradient VE40 VR Dead space estimates Compliance Driving pressure

Abbreviations: BMI: Body mass index, A-a: Aleolar-arterial, VE40: corrected minute ventilation, VR: ventilatory ratio

Table 4. Demographics of excluded datasets and Initial cohort

Characteristic	Excluded Datasets n=2410	Initial Cohort n=1907
Age, mean (std)	51 (17)	53 (16)
Female	47%	50%
Race		
White	70%	77%
Black	18%	16%
Other/missing	12%	7%

Abbreviations: std: standard deviation

Table 5. Baseline Characteristics of Initial Cohort

Characteristic, no. (%)	All Participants n=1907 (%)
Age, Mean (std)	52.87 (16.22)
18-29 yrs	186 (9.75)
30-39 yrs	205 (10.75)
40-49 yrs	375 (19.66)
50-59 yrs	497 (26.06)
60-69 yrs	344 (18.04)
70+ yrs	300 (15.73)
Female Gender	943 (49.45)
Race	
White	1472 (77.19)
Black	296 (15.52)
Other	71 (3.72)
Missing	68 (3.57)
Place of Residence	
Home Independently	1545 (81.02)
Home with help	193 (10.12)
Home w professional help	27 (1.42)
Intermediate Care or rehab	33 (1.73)
Skilled nursing facility	75 (3.93)
Other	33 (1.73)
Missing	1 (0.05)
Co-morbid conditions	
Chronic lung disease Missing	260 (13.63) 1 (0.05)
On chronic dialysis	55 (2.88)
Cirrhosis Missing	95 (4.98) 2 (0.1)
Solid Tumor with mets	53 (2.78)
Liquid Tumor	89 (4.67)
Diabetes Missing	486 (25.49) 1 (0.05)
Hypertension Missing	886 (46.46) 2 (0.1)
Previous MI Missing	102 (5.35) 1 (0.05)
Congestive Heart Failure Missing	119 (6.24) 1 (0.05)
Prior stroke with sequelae Missing	58 (3.04) 1 (0.05)

Immunosuppressed	239 (12.53)
Health Behaviors	
Tobacco Use	
Ever Smoker Missing	1002 (52.54) 183 (9.6)
Current Smoker Missing	568 (29.79) 906 (47.51)
Alcohol Frequency	
Never	918 (48.14)
Monthly	282 (14.79)
2-4 times/month	127 (6.66)
2-3 times/wk	109 (5.72)
4+ times/wk	282 (14.79)
Missing	189 (9.91)
Height, cm Mean (std) Missing	169.02 (10.9) 1 (0.05)
Weight, kg Median (IQR) Missing	82 (68, 100) 6 (0.31)
BMI, kg/m^2 Median (IQR)	28.61 (23.85, 34.48)
BMI < 18	60 (3.15)
BMI 18-24	531 (27.84)
BMI 25-30	477 (25.01)
BMI > 30	832 (43.63)
Missing	7 (0.37)
Ideal Body weight, kg Median (IQR) Missing	62 (55, 72) 2 (0.1)
APACHE III score Mean (std) Missing	92.24 (27.75) 62 (3.25)
No. of organ failure Median (IQR)	1 (1, 2)
Cardiac Failure	554 (29.05)
Hepatic Failure	1624 (85.16)
Renal Failure	1439 (75.46)
CNS Failure	209 (10.96)
Coagulopathy	1560 (81.8)
Medical ICU	1604 (84.11)
Baseline vasopressors	767 (40.22)
Baseline Shock	825 (43.26)
ARDS Risk Factor	
Pneumonia	1211 (63.5)
Sepsis	339 (17.78)
Trauma	64 (3.36)

Aspiration	186 (9.75)
Multiple Transfusions	25 (1.31)
Other/not classified	82 (4.3)
ARDS Severity by Berlin Criteria	
Mild	301 (15.78)
Moderate	824 (43.21)
Severe	549 (28.79
Missing	233 (12.22)
PaO2/FiO2 Median (IQR)	128 (87, 179)
Pulmonary Worsening 48 hours Missing	1121 (58.78) 59 (3.09)
28-Day Mortality	403 (21.13)

Abbreviations: std: standard deviation, IQR: interquartile range. All values given as no.(%) unless otherwise stated. APACHE: Acute Physiology and Chronic Health Evaluation, ARDS: Acute Respiratory Distress syndrome

Table 6. Baseline Pulmonary Characteristics of Initial Cohort

Physiologic Parameter, Median (IQR)	All Participants n=1907
Oxygenation Index **Missing, no. (%)**	9.42 (5.87, 14.59) 213 (11.17%)
Mean Airway Pressure, cmH2O	15 (12, 18)
PaO2, mmHg	83 (70, 106)
FiO2	0.5 (0.4, 0.7)
Oxygenation Saturation Index **Missing, no. (%)**	8.16 (5.42, 12.27) 190 (9.96%)
SpO2 %	96 (94, 98)
A-a Gradient mmHg **Missing, no. (%)**	238.57 (163.32, 358.3) 40 (2.10%)
PAO2 mmHg	318.07 (248.71, 449.3)
a/A Ratio	0.26 (0.18, 0.37) 40 (2.10%)
Ventilatory Ratio **Missing, no. (%)**	1.76 (1.4, 2.19) 112 (5.87%)
Minute Ventilation L/min	10.6 (8.6, 12.9)
PaCO2 mmHg	38 (34, 45)
Predicted VE (IBW x 100) L/min	6.2 (5.5, 7.2)
VE40 L/min **Missing, no. (%)**	10.15 (8.19, 12.74) 110 (5.77%)
Siddiki Dead Space **Missing, no. (%)**	0.38 (0.14, 0.55) 117 (6.14%)

HB_sid basal metabolic rate kcal/d	1449.22 (1218.34, 1789.54)
hf correction factor	1.6 (1.6, 2.15)
Harris-Benedict Dead Space	0.6 (0.49, 0.68)
Missing, no. (%)	117 (6.14%)
HB basal metabolic rate kcal/d	1596.3 (1389.1, 1877.28)
Pulmonary Compliance ml/cmH2O	30.63 (23.13, 40)
Missing, no. (%)	708 (37.13%)
Tidal volume mL	410 (355, 474)
Plateau Pressure cmH2O Mean (std)	23.65 (6.07)
PEEP cmH2O	10 (5, 12)
Driving Pressure cmH2O	14 (11, 18)
Missing, no. (%)	648 (33.98%)

Abbreviations: IQR: interquartile range, PaO2: partial pressure of oxygen in the blood, FiO2: fraction of inspired oxygen, SpO2: saturation of oxygen A: alveolar, a: arterial, VE: Minute ventilation HB: Harris-Benedict, hf: hypermetabolic factor, PEEP: Positive end-expiratory pressure.
All values given as median (IQR) unless otherwise stated

Table 7. Baseline Characteristics of Derivation Cohort

Characteristic	Pulmonary Improvement n=301 (39.5%)	Pulmonary Worsening n=461 (60.5%)	p-value[†]
Age, Mean (std)	52.12 (17.11)	54.14 (15.98)	0.097
18-29 yrs	13.29	8.24	0.0469
30-39 yrs	9.97	10.20	
40-49 yrs	16.28	19.52	
50-59 yrs	29.90	24.95	
60-69 yrs	13.95	20.17	
70+ yrs	16.61	16.92	
Female Gender	49.17	50.76	0.6679
Race			0.2336
White	77.08	77.22	
Black	17.61	14.1	
Other	2.66	4.56	
Missing	2.66	4.12	
Place of Residence			0.7653
Home Independently	83.39	80.04	
Home with help	8.64	11.28	
Home w professional help	1.33	1.30	
Intermediate Care or rehab	0.66	1.30	

Skilled nursing facility	3.99	4.56	
Other	1.99	1.52	
Co-morbid conditions			
Chronic lung disease	9.30	13.88	0.0578
On chronic dialysis	3.32	3.25	0.9586
Cirrhosis	4.31	4.12	0.8943
Solid Tumor with mets	2.66	3.25	0.6383
Liquid Tumor	2.66	6.07	0.0298
Diabetes	29.24	26.03	0.3315
Hypertension	47.51	44.47	0.4102
Previous MI	6.64	6.72	0.9655
Congestive Heart Failure	4.98	6.29	0.4495
Prior stroke with sequelae	2.33	2.60	0.8102
Immunosuppressed	9.63	11.93	0.3225
Health Behaviors			
Tobacco Use			
Ever Smoker Missing	53.82 10.63	51.19 11.06	0.4909
Current Smoker Missing	32.56 47.18	29.07 48.81	0.3364
Alcohol Frequency			0.5268
Never	43.85	47.07	
Monthly	14.95	15.84	
2-4 times/month	6.98	4.34	
2-3 times/wk	5.98	5.21	
4+ times/wk	16.94	15.84	
Missing	11.30	11.71	
Height, cm Mean (std)	169.47 (11.46)	169.13 (11.37)	0.6846
Weight, kg Median (IQR)	81 (65, 96)	79 (66, 98)	0.7577
BMI, kg/m^2 Median (IQR)	27.50 (23.07, 33.35)	27.56 (23.45, 33.95)	0.5381
BMI < 18	2.99	3.69	0.9083
BMI 18-24	33.22	31.89	
BMI 25-30	23.92	25.38	
BMI > 30	39.87	39.05	
Ideal Body weight, kg Median (IQR) Missing	62 (55, 73)	62 (55, 72)	0.6896
APACHE III score Mean (std) Missing	91.87 (28.22)	93.95 (28.14)	0.3228

No. of organ failure Median (IQR)	1 (1, 2)	1 (1, 2)	0.2066
Cardiac Failure	26.91	29.50	0.4385
Hepatic Failure	89.37	83.51	0.0235
Renal Failure	80.73	76.57	0.1740
CNS Failure	11.30	12.36	0.6565
Coagulopathy	84.72	81.34	0.2291
Medical ICU	83.72	79.83	0.1771
Baseline vasopressors	41.20	42.95	0.6318
Baseline Shock	45.18	46.85	0.6509
ARDS Risk Factor			0.0615
Pneumonia	62.46	63.56	
Sepsis	14.29	18.22	
Trauma	2.33	4.34	
Aspiration	14.62	8.46	
Multiple Transfusions	1.00	0.65	
Other/not classified	5.32	4.77	
ARDS Severity by Berlin Criteria			0.0130
Mild	10.63	18.22	
Moderate	47.18	46.85	
Severe	33.22	27.77	
Missing	8.97	7.16	
PaO2/FiO2 Median (IQR)	124 (80, 173)	135 (93, 185)	0.0144
28-Day Mortality	9.3	25.38	<0.0005

Abbreviations: std: standard deviation, IQR: interquartile range. All values given as no.(%) unless otherwise stated.
APACHE: Acute Physiology and Chronic Health Evaluation, ARDS: Acute Respiratory Distress syndrome
† p-value for t-test, rank-sum, and chi sq for difference in means, median, and proportions respectively

Table 8. Baseline Pulmonary Characteristics of Derivation Cohort

Physiologic Parameter, Median (IQR)	Pulmonary Improvement n=301	Pulmonary Worsening n=461	p-value†
Oxygenation Index	10.38 (6.77, 16)	8.77 (5.52, 13.10)	0.0009
Mean Airway Pressure, cmH2O	16 (13, 18)	14 (11, 17)	<0.0001
PaO2, mmHg	87 (72, 111)	84 (70, 108)	0.1859
FiO2	0.6 (0.5, 0.7)	0.5 (0.4, 0.7)	<0.0001
Oxygenation Saturation Index	9.38 (6.38, 13.8)	7.65 (5.2, 11.05)	<0.0001
SpO2 %	97 (94, 99)	96 (94, 98)	0.0017

A-a Gradient mmHg	269.55 (179.21, 385.05)	231.32 (159.21, 338.05)	0.0004
PAO2 mmHg	373.68 (294.32, 461.8)	314.32 (246.21, 425.55)	<0.0001
a/A Ratio	0.25 (0.18, 0.35)	0.27 (0.19, 0.38)	0.0476
Ventilatory Ratio	1.7 (1.39, 2.24)	1.8 (1.42, 2.19)	0.4949
Minute Ventilation L/min	10.8 (8.8, 12.8)	10.5 (8.8, 12.8)	0.7484
PaCO2 mmHg	38 (33, 44)	38 (33, 44)	0.9765
Predicted VE (IBW x 100) L/min	6.2 (5.5, 7.3)	6.2 (5.5, 7.2)	0.6896
VE40 L/min	10.15 (8.17, 12.6)	10.4 (8.19, 12.92)	0.6767
Siddiki Dead Space	0.4 (0.16, 0.56)	0.4 (0.16, 0.56)	0.9876
HB_sid basal metabolic rate kcal/d	1403.2 (1195.92, 1800.11)	1407.63 (1184.79, 1767.63)	0.5896
hf correction factor	1.6 (1.6, 2.15)	1.6 (1.6, 2.03)	0.9147
Harris-Benedict Dead Space	0.61 (0.52, 0.68)	0.62 (0.51, 0.69)	0.7903
HB basal metabolic rate kcal/d	1581.51 (1384.42, 1834.17)	1551.22 (1366.01, 1834.43)	0.6596
Pulmonary Compliance ml/cmH2O	30.77 (24, 41.25)	30.77 (22.78, 39.09)	0.5407
Tidal volume mL	410 (360, 468)	420 (350, 480)	0.533
Plateau Pressure cmH2O Mean (std)	24.39 (6.1)	23.08 (5.71)	0.0028
PEEP cmH2O	10 (8, 12)	8 (5, 10)	<0.0001
Driving Pressure cmH2O	14 (10, 16)	14 (11, 18)	0.0947

Abbreviations: IQR: interquartile range, PaO2: partial pressure of oxygen in the blood, FiO2: fraction of inspired oxygen, SpO2: saturation of oxygen A: alveolar, a: arterial, VE: Minute ventilation HB: Harris-Benedict, hf: hypermetabolic factor, PEEP: Positive end-expiratory pressure.
All values given as median (IQR) unless otherwise stated
Ŧ p-value for t-test and rank-sum for difference in means and medians respectively

Table 9. How Participants qualified for primary outcome in the derivation set

Outcome component, no. (%)	Pulmonary worsening in Derivation cohort N = 461 (60.5%)
Early Mortality	19 (4.1)
PEEP and FiO2	216 (46.9)
PEEP	151 (32.8)
FiO2	74 (16.1)
PEEP w/ missing FiO2	1 (0.2)
FiO2 w/ missing PEEP	0 (0)

Abbreviations: PEEP: positive end-expiratory pressure, FiO2: fraction of inspired oxygen

Table 10. Final Logistic Model Parameters

Covariate	**β estimate (SE)**	**Odds Ratio (95% CI)**	**P-value†**
Oxygenation Saturation Index	-0.0678 (0.0141)	0.934 (0.909, 0.961)	<0.0001
Driving Pressure	0.0376 (0.0154)	1.038 (1.008, 1.070)	0.0143
Hematologic Malignancy	0.8395 (0.4133)	2.315 (1.030, 5.204)	0.0422
Chronic Pulmonary Disease	0.4580 (0.2452)	1.581 (0.978, 2.557)	0.0618
Hepatic Failure	-0.5855 (0.2313)	0.557 (0.354, 0.876)	0.0113

Abbreviations: CI: confidence interval
† Wald-chi sq test p-value

Table 11. Iterations of Model Refinement

Model	AUC	-2 Log L
Base (derived from backward selection with p < 0.1)	**0.6432**	**981.245**
OSI squared	0.6361	988.307
OSI cubed	0.6255	994.519
Ln (OSI)	0.6421	979.140
DP squared	0.6429	981.943
DP cubed	0.6409	983.869
Ln(DP)	0.6417	982.615
BMI and chronic pulmonary dx interaction with OSI and DP	0.6462	971.542
BMI interaction with OSI and DP	0.6482	972.673
Age back	0.6460	980.301
Gender back	0.6433	981.180
BMI back	0.6439	980.760
ARDS cause back	0.6545	972.634
Categorized OSI and driving pressure using LOESS to find cutpoints	0.6480	974.105

Abbreviations: AUC: Area under receiver operating curve, OSI: oxygen saturation index, Ln: natural log, DP: Driving pressure, BMI: body mass index, dx: disease, ARDS: Acute Respiratory Distress Syndrome

Table 12. Classification Table of the Final Model

Probability	**Accuracy**	**Sensitivity**	**Specificity**
0.48	62.3	90.7	18.9
0.52	63.5	87.2	27.2
0.56	61.3	76.8	37.5
0.58	60.8	70.7	45.5
0.64	55.6	43.8	73.8
0.66	53.9	36	81.4
0.70	48.4	21.0	90.4

FIGURES

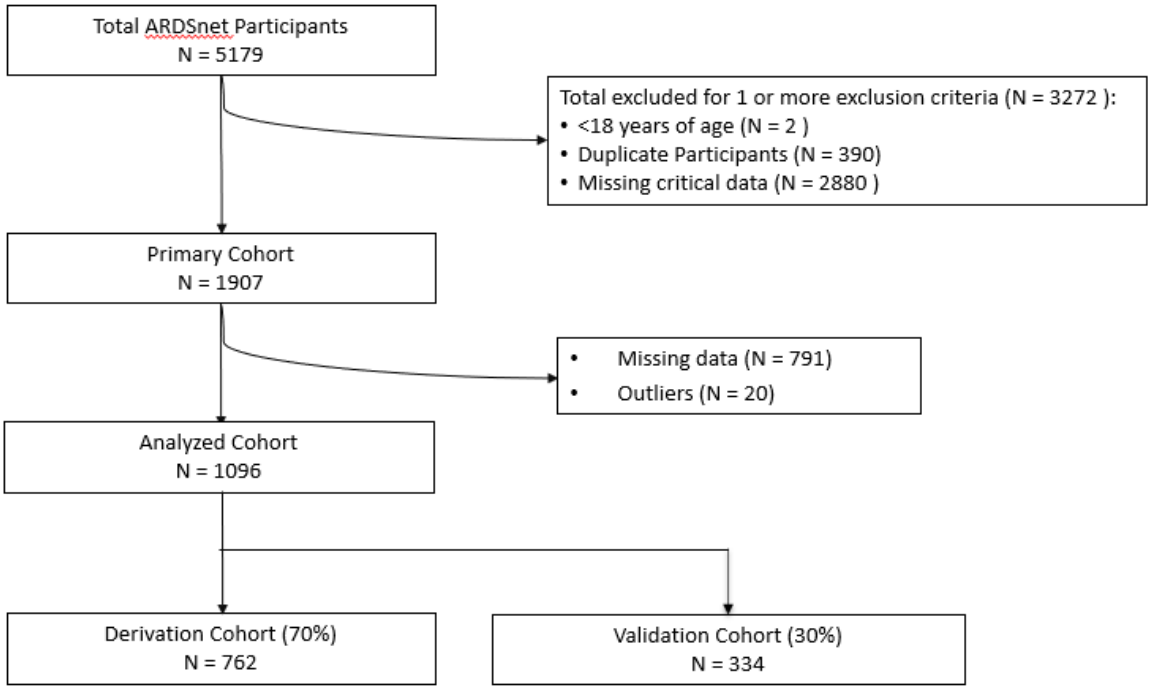

Figure 1. Flow Diagram of Participant Selection into Final Cohorts

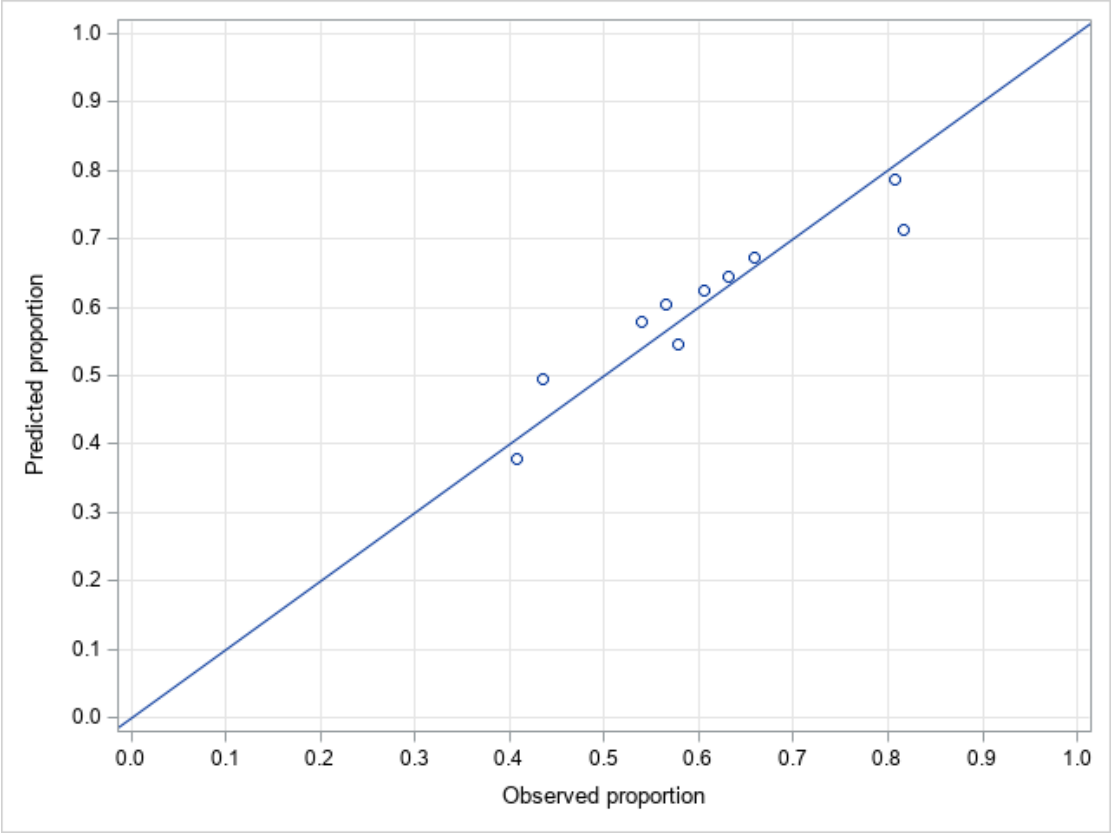

Figure 2. Calibration Plot of Final Logistic Model

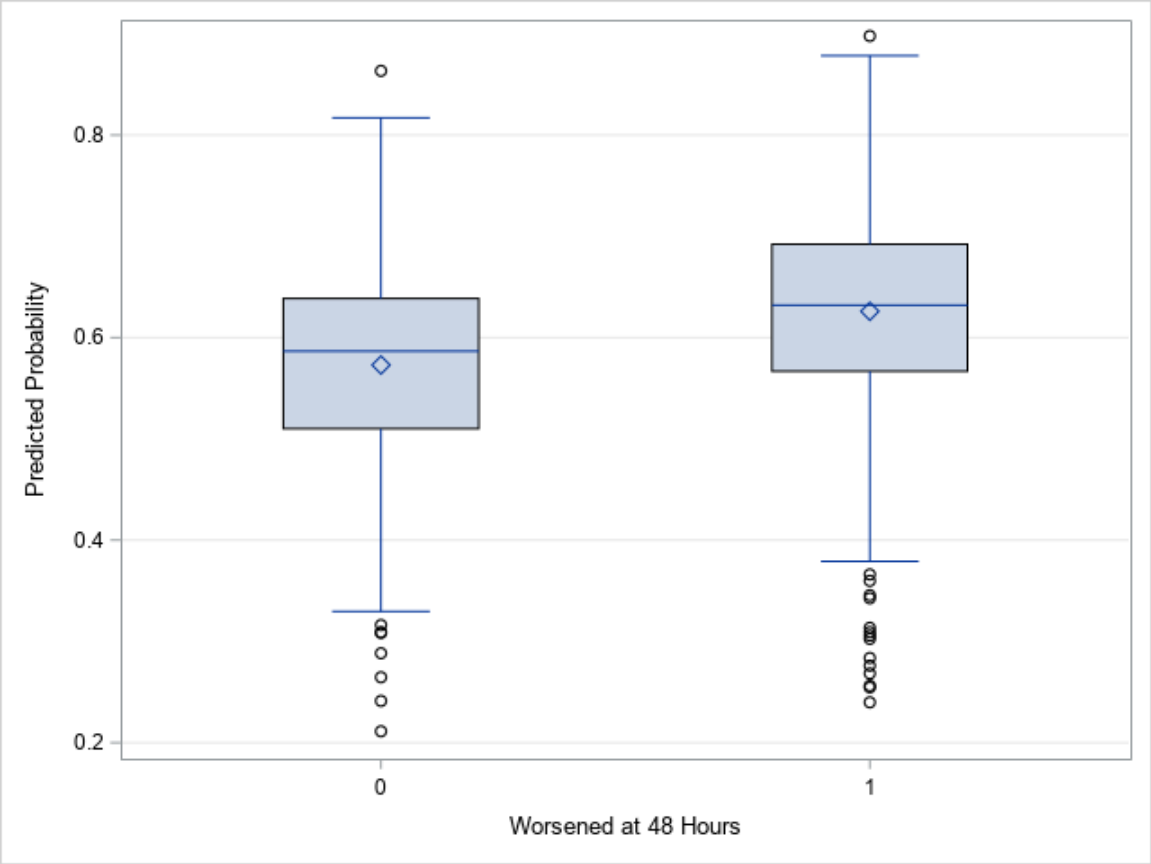

Figure 3. Box and Whisker Plot of Final Logistic Model
Line through the box indicates median, boundaries of the box indicate the interquartile range, the diamond indicates the mean, circles represent outliers.
1 = experienced outcome, 0 = did not experience outcome

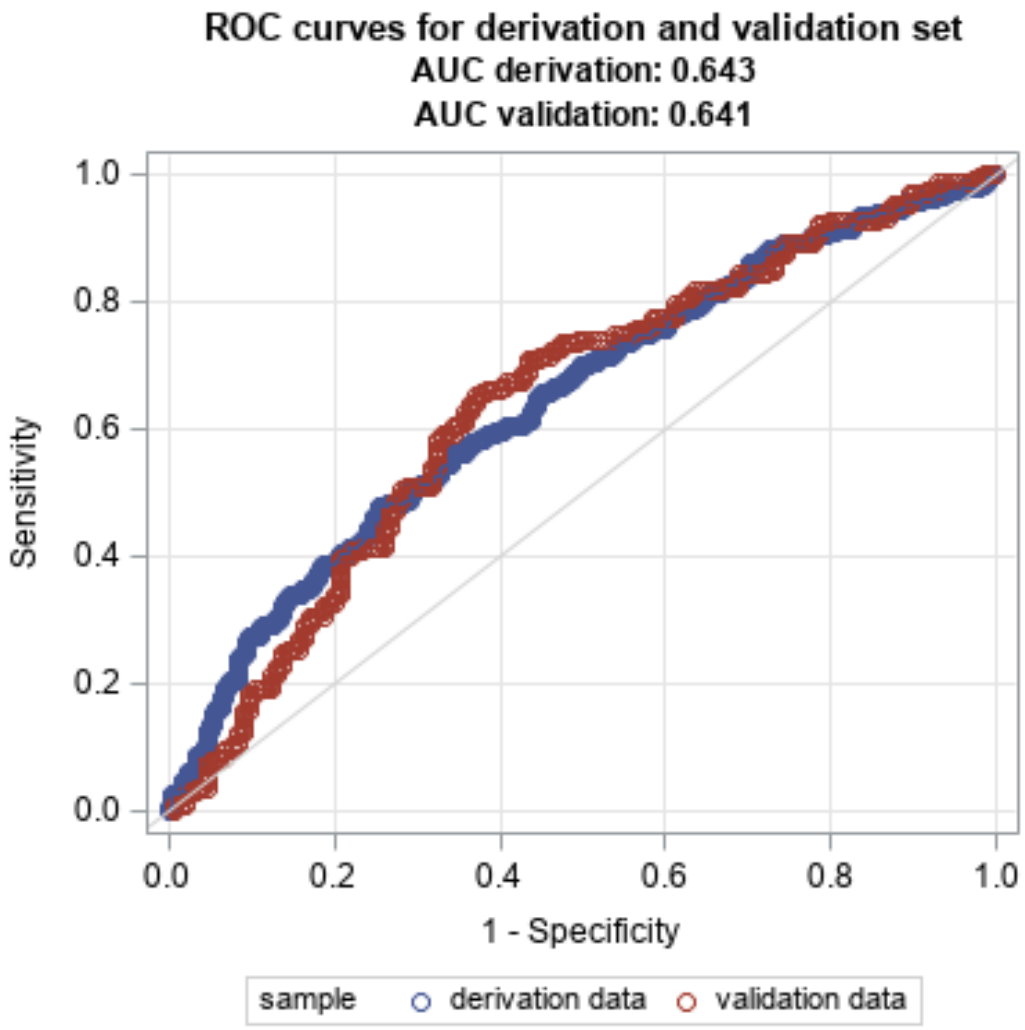

Figure 4. Receiver Operating Curves of Final Model