

Predicting Mortality in Children With Pediatric Acute Respiratory Distress Syndrome: A Pediatric Acute Respiratory Distress Syndrome Incidence and Epidemiology Study

Nadir Yehya, MD¹; Michael O. Harhay, PhD^{2,3}; Margaret J. Klein, MS⁴; Steven L. Shein, MD⁵; Byron E. Piñeres-Olave, MD⁶; Ledys Izquierdo, MD⁷; Anil Sapru, MD⁸; Guillaume Emeriaud, MD, PhD⁹; Philip C. Spinella, MD, FCCM¹⁰; Heidi R. Flori, MD¹¹; Mary K. Dahmer, PhD¹¹; Aline B. Maddux, MD, MSCS¹²; Yolanda M. Lopez-Fernandez, MD¹³; Bereketeab Haileselassie, MD, MHS¹⁴; Deyin Doreen Hsing, MD¹⁵; Ranjit S. Chima, MD¹⁶; Amanda B. Hassinger, MD, MS¹⁷; Stacey L. Valentine, MD, MPH¹⁸; Courtney M. Rowan, MD¹⁹; Martin C. J. Kneyber, MD, PhD, FCCM²⁰; Lincoln S. Smith, MD²¹; Robinder G. Khemani, MD, MsCI⁴; Neal J. Thomas, MD, MSc²²; on behalf of the Pediatric Acute Respiratory Distress Syndrome Incidence and Epidemiology (PARDIE) V1 Investigators and the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) Network

¹Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia, PA.

²Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA.

³Palliative and Advanced Illness Research (PAIR) Center, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA.

⁴Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Los Angeles and University of Southern California, Los Angeles, CA.

⁵Department of Pediatrics, Division of Pediatric Critical Care Medicine, Rainbow Babies and Children's Hospital and Case Western Reserve University, Cleveland, OH.

⁶Pediatric Critical Care Unit, Hospital Pablo Tobón Uribe, Medellín, Colombia.

⁷Department of Pediatrics, Hospital Militar Central, Bogotá, Colombia.

⁸Department of Pediatrics, Division of Critical Care, University of California Los Angeles Mattel Children's Hospital, Los Angeles, CA.

⁹Department of Pediatrics, Centre Hospitalier Universitaire Sainte-Justine and Université de Montréal, Montreal, QC, Canada.

¹⁰Department of Pediatrics, Division of Critical Care, Washington University in St. Louis, St. Louis, MO.

¹¹Department of Pediatrics, Division of Pediatric Critical Care Medicine, Mott Children's Hospital and University of Michigan, Ann Arbor, MI.

¹²Department of Pediatrics, Children's Hospital Colorado and University of Colorado, Aurora, CO.

¹³Department of Pediatrics, Cruces University Hospital, Bizkaia, Basque Country, Spain.

¹⁴Department of Pediatrics, Division of Pediatric Critical Care, Stanford University, Palo Alto, CA.

¹⁵Department of Pediatrics, New York-Presbyterian Hospital and Weill Cornell Medical College, New York, NY.

¹⁶Department of Pediatrics, Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center and University of Cincinnati, Cincinnati, OH.

¹⁷Department of Pediatrics, Division of Pediatric Critical Care, Oishei Children's Hospital and University of Buffalo, Buffalo, NY.

¹⁸Department of Pediatrics, Division of Pediatric Critical Care, University of Massachusetts Children's Medical Center, Worcester, MA.

¹⁹Department of Pediatrics, Division of Critical Care, Riley Hospital for Children and Indiana University, Indianapolis, IN.

²⁰Department of Pediatrics, Division of Pediatric Critical Care Medicine, Beatrix Children's Hospital and University of Groningen, Groningen, The Netherlands.

²¹Department of Pediatrics, Division of Pediatric Critical Care Medicine, Seattle Children's Hospital and University of Washington, Seattle, WA.

²²Department of Pediatrics and Public Health Science, Division of Pediatric Critical Care Medicine, Penn State Hershey Children's Hospital, Hershey, PA.

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the Society of Critical Care Medicine and Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1097/CCM.0000000000004345

Objectives: Pediatric acute respiratory distress syndrome is heterogeneous, with a paucity of risk stratification tools to assist with trial design. We aimed to develop and validate mortality prediction

models for patients with pediatric acute respiratory distress syndrome.

Design: Leveraging additional data collection from a preplanned ancillary study (Version 1) of the multinational Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology study, we identified predictors of mortality. Separate models were built for the entire Version 1 cohort, for the cohort excluding neurologic deaths, for intubated subjects, and for intubated subjects excluding neurologic deaths. Models were externally validated in a cohort of intubated pediatric acute respiratory distress syndrome patients from the Children's Hospital of Philadelphia.

Setting: The derivation cohort represented 100 centers worldwide; the validation cohort was from Children's Hospital of Philadelphia.

Patients: There were 624 and 640 subjects in the derivation and validation cohorts, respectively.

Interventions: None.

Measurements and Main Results: The model for the full cohort included immunocompromised status, Pediatric Logistic Organ Dysfunction 2 score, day 0 vasopressor-inotrope score and fluid balance, and $\text{PaO}_2/\text{FiO}_2$ 6 hours after pediatric acute respiratory distress syndrome onset. This model had good discrimination (area under the receiver operating characteristic curve 0.82), calibration, and internal validation. Models excluding neurologic deaths, for intubated subjects, and for intubated subjects excluding neurologic deaths also demonstrated good discrimination (all area under the receiver operating characteristic curve ≥ 0.84) and calibration. In the validation cohort, models for intubated pediatric acute respiratory distress syndrome (including and excluding neurologic deaths) had excellent discrimination (both area under the receiver operating characteristic curve ≥ 0.85), but poor calibration. After revision, the model for all intubated subjects remained miscalibrated, whereas the model excluding neurologic deaths showed perfect calibration. Mortality models also stratified ventilator-free days at 28 days in both derivation and validation cohorts.

Conclusions: We describe predictive models for mortality in pediatric acute respiratory distress syndrome using readily available variables from day 0 of pediatric acute respiratory distress syndrome which outperform severity of illness scores and which demonstrate utility for composite outcomes such as ventilator-free days. Models can assist with risk stratification for clinical trials. (*Crit Care Med* 2020; 48:e514–e522)

Key Words: mortality; pediatric acute respiratory distress syndrome; prediction; risk stratification; ventilator-free days

Acute respiratory distress syndrome (ARDS) is characterized by acute hypoxemic respiratory failure from noncardiogenic pulmonary edema. Despite inclusion of pediatric subjects in the initial description of ARDS (1), neither the 1994 American-European Consensus Conference (2) nor the 2012 Berlin revised definitions of ARDS (3) addressed pediatric considerations. To inform study design in children with ARDS, the Pediatric Acute Lung Injury Consensus Conference (PALICC) developed a specific definition for pediatric ARDS (PARDS) in 2015 (4). Notable differences in the

PALICC definition include use of oxygenation index (OI), rather than $\text{PaO}_2/\text{FiO}_2$, for severity stratification; explicit use of alternative stratification using peripheral oxygen saturation (SpO_2) (oxygen saturation index [OSI]); and inclusion of unilateral, in addition to bilateral, infiltrates on chest radiograph. Both the Berlin and PALICC definitions highlighted the need to better stratify risk, with PALICC specifically recommending the development of prognostic scores to stratify mortality (4, 5). Appropriate risk stratification would permit prognostic enrichment (restriction of trial eligibility to those at highest risk of poor outcome) for clinical trials (6), as well as allow testing for heterogeneity of treatment effect (7).

Recently, the Pediatric ARDS Incidence and Epidemiology (PARDIE) study was completed to assess the utility of the PALICC definition, identifying 708 subjects from 145 PICUs worldwide (8). PARDIE identified three-fold more cases of PARDS using PALICC criteria, compared with Berlin, with improved mortality discrimination with PALICC severity categories. In a preplanned ancillary study (dubbed Version 1 [V1]), additional data were collected to assess predictors of mortality. Herein, we present the development and validation of mortality prediction models in PARDS. We hypothesized that a combination of clinical variables could reliably estimate the probability of PICU mortality.

MATERIALS AND METHODS

The full Methods are available in the **Data Supplement** (Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). The study was designed and reported consistent with the recommendations of the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement (**Appendix 1**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). The Children's Hospital of Los Angeles (CHLA, Los Angeles, CA) was the clinical and data coordinating center. The protocol was first approved by the CHLA Institutional Review Board (IRB). PARDIE sites obtained ethical approval from their local IRB or relied on the central CHLA IRB.

Derivation (PARDIE V1) Cohort

PARDIE consisted of 708 subjects from 145 PICUs in 27 predominantly high-income countries (8), prospectively screened over 5 days during 10 nonconsecutive weeks between May 2016 and June 2017, with each site deciding a priori whether to participate in the ancillary studies. V1 included additional data from 100 centers (**Appendix 2**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). Patients were eligible for PARDIE if they newly met PALICC PARDS criteria. We collected data for the first 3 days after PARDS diagnosis, including demographics, center characteristics, geoeconomic categorization, oxygenation, ventilatory support, radiographs, severity of illness measured using the Pediatric Index of Mortality (PIM) 3 and Pediatric Risk of Mortality (PRISM) IV scores, and comorbidities. Additional data collected for V1 included daily (calendar day) organ failure (Pediatric Logistic Organ Dysfunction [PELOD] 2 score), vasopressor requirement, fluid balance, and use of ancillary therapies.

External Validation Cohort

We externally validated the predictive models using a prospective cohort of intubated children meeting Berlin ARDS criteria from the PICU of the Children's Hospital of Philadelphia (CHOP), a 60-bed PICU in a quaternary North American children's hospital, between July 2011 and June 2018. As CHOP was a participant in PARDIE, overlapping subjects were excluded.

Definitions and Outcomes

The primary outcome was PICU mortality. Secondary outcomes include duration of ventilation in survivors and ventilator-free days (VFDs) at 28 days. Oxygenation was measured using P_{aO_2}/F_{iO_2} and S_{pO_2}/F_{iO_2} in all subjects, and OI and OSI in intubated subjects (9, 10). For all analyses, noninvasive measures were converted to invasive equivalents using published equations (9). Vasopressor-inotrope score was dopamine ($\mu\text{g}/\text{kg}/\text{min}$) \times 1 + dobutamine ($\mu\text{g}/\text{kg}/\text{min}$) \times 1 + epinephrine ($\mu\text{g}/\text{kg}/\text{min}$) \times 100 + norepinephrine ($\mu\text{g}/\text{kg}/\text{min}$) \times 100 + milrinone ($\mu\text{g}/\text{kg}/\text{min}$) \times 10 (11). The designation "immunocompromised" required presence of an oncologic diagnosis, immunodeficiency, stem cell or organ transplant, or a rheumatologic or inflammatory condition receiving immunosuppression (12, 13). Countries were grouped by geographical region and economic status using 2016 World Bank classifications (14). A single cause of death was assigned by site investigators: hypoxemia, refractory shock, multisystem organ failure (MSOF), brain death, other neurologic cause, or other.

Development of a Model for Mortality Prediction

Our primary aim was to construct a parsimonious model of clinical variables on day of PARDS onset (day 0) associated with PICU mortality for use in risk prediction. We did this in two steps: penalized regression followed by variable reduction using the Bayesian information criterion (BIC), detailed in the Data Supplement (Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). Internal validation was evaluated by 10-fold cross-validation and assessment of model performance in prespecified subgroups. Calibration and fit were assessed using the calibration belt (15). Discrimination for PICU mortality was assessed by calculating area under the receiver operating characteristic (AUROC) curve.

The primary model was developed using the entire V1 cohort, which included intubated and noninvasively ventilated subjects with PARDS, and all-cause mortality. Three subgroup models were built. First, as subjects dying of a neurologic etiology may have different predictors of mortality than those dying of shock, MSOF, or hypoxemia (16), we repeated the analysis excluding those who died primarily due to a neurologic cause. Second, we repeated the analysis in patients invasively ventilated within 6 hours of PARDS diagnosis. Third, we repeated the analysis in invasively ventilated subjects excluding those who died from a neurologic cause.

External Validation of the Model

Models were validated using a cohort of children with Berlin ARDS from CHOP. As all CHOP subjects were intubated, we

only assessed models for invasively ventilated subjects. Calibration, fit, and discrimination were reported. Since the model was derived from a multicenter cohort, we reasoned that if calibration was poor, the model would be revised in this cohort by reestimation of the coefficients and intercept (17).

Utility of the Mortality Model to Stratify Ventilator-Free Days

We assessed whether models developed for mortality were calibrated for VFDs in both derivation and validation cohorts. Subjects were split into quartiles of predicted mortality for each of the models. For each quartile, VFDs were modeled as a competing risk, treating discontinuation of ventilation as the primary outcome, and death as a competing event. Outcomes were censored after 28 days, making this equivalent to VFDs at 28 days (18). Models were constructed for all patients (invasive and noninvasive VFDs) and for intubated patients (invasive VFDs) in the derivation cohort and limited to invasive VFDs for the validation cohort. Noninvasive ventilation was defined as oro-nasal mask continuous positive airway pressure or bilevel positive airway pressure. High-flow oxygen was not counted as noninvasive support.

Development of a Model for Identifying Predictors of Ventilator Duration

As mortality in PARDS is low, length of ventilation in survivors is an important contributor to VFDs. Therefore, we separately constructed models to identify predictors of total (invasive and noninvasive) and invasive ventilator duration in survivors. We modeled ventilator duration as time to event analyses using Cox regression with clustering by site, censored at 28 days, with variable selection based on BIC optimization, similar to the mortality models.

RESULTS

Description of the V1 Cohort

Description of the V1 cohort ($n = 624$, 108 deaths) stratified by PICU survival status is presented in **Table 1**, and more completely in **Supplementary Table 1** (Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). Nonsurvivors had greater severity of illness scores, more comorbidities, and higher vasopressor scores and fluid balance on the day of PARDS diagnosis (day 0). Nonsurvivors were more likely to have worse oxygenation, irrespective of whether invasively ventilated or not. Finally, nonsurvivors were more likely to come from World Bank-designated middle-income countries. Of the 108 nonsurvivors, 28 (26%) had a primary neurologic etiology as a cause of death.

Predictive Models for PICU Mortality from the V1 Cohort

Table 2 shows the final model for PICU mortality in the entire cohort. PELOD 2 on day of PARDS diagnosis (day 0), vasopressor score and fluid balance (mL/kg) on day 0 of PARDS, immunocompromised status, and P_{aO_2}/F_{iO_2} 6 hours after PARDS

TABLE 1. Description of the Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology Version 1 Cohort Stratified by Mortality

Variable	Whole Cohort (n = 624)	Survivors (n = 516)	Nonsurvivors (n = 108)	p
Demographics				
Age (yr) (n = 619), mean ± SD	6.1 ± 6.4	6.0 ± 6.3	6.5 ± 6.6	0.513
Female/male (%) (n = 623)	247/376 (40/60)	205/310 (40/60)	42/66 (39/61)	0.914
Nonwhite race (%) (n = 619)	243 (39)	193 (38)	50 (46)	0.105
Hispanic ethnicity (%) (n = 621)	139 (22)	104 (20)	35 (32)	0.008
Admission source (%) (n = 623)				
Emergency department	286 (46)	245 (48)	41 (38)	0.018
Inpatient floor	219 (35)	168 (33)	51 (48)	
Other	118 (19)	102 (20)	16 (15)	
Severity of illness, mean ± SD				
Pediatric Index of Mortality 3 (admission)	-3.6 ± 2.1	-3.9 ± 1.7	-2.1 ± 2.7	< 0.001
Pediatric Risk of Mortality IV (admission) (n = 622)	9.0 ± 8.3	7.7 ± 7.1	15.0 ± 10.5	< 0.001
Pediatric Logistic Organ Dysfunction 2 (PARDS onset) (n = 621)	5.4 ± 3.3	4.9 ± 2.8	7.9 ± 4.3	< 0.001
Vasopressor-inotrope score day 0 (n = 619), mean ± SD	11.6 ± 29.6	7.4 ± 21.4	32.1 ± 48.9	< 0.001
Fluid balance day 0 (mL/kg) (n = 609), mean ± SD	26.6 ± 43.4	23.7 ± 39.3	41.2 ± 57.1	< 0.001
Preexisting comorbidities (%)				
None	233 (37)	208 (40)	25 (23)	0.001
Immunocompromised	82 (13)	42 (8)	40 (37)	< 0.001
PARDS etiology (%)				
Pneumonia	392 (63)	341 (66)	51 (47)	< 0.001
Nonpulmonary sepsis	121 (19)	86 (17)	35 (32)	
Other	111 (18)	89 (17)	22 (20)	
Pao ₂ /Fio ₂ , mean ± SD				
PARDS diagnosis ^a	142 ± 83	148 ± 84	115 ± 69	< 0.001
6 hr	177 ± 108	185 ± 109	138 ± 91	< 0.001
Oxygenation index in intubated subjects, mean ± SD				
PARDS diagnosis (n = 485) ^a	14.1 ± 11.1	12.8 ± 9.8	20.5 ± 13.9	< 0.001
6 hr (n = 531) ^a	12.5 ± 11.2	10.7 ± 8.8	20.5 ± 16.5	< 0.001
Pediatric Acute Lung Injury Consensus Conference categories (%)				
Noninvasive	139 (22)	117 (23)	22 (20)	< 0.001
Mild	200 (32)	175 (34)	25 (23)	
Moderate	133 (21)	121 (23)	12 (11)	
Severe	152 (24)	103 (20)	49 (45)	
Goeconomic status (%)				
High income: North America	413 (66)	351 (68)	62 (57)	0.001
High income: Europe	91 (15)	77 (15)	14 (13)	
High income: rest of world	36 (6)	32 (6)	4 (4)	
Middle income	84 (13)	56 (11)	28 (26)	

PARDS = pediatric acute respiratory distress syndrome.

^aInvasive measures of oxygenation (Pao₂/Fio₂ and oxygenation index) include values derived from noninvasive (peripheral oxygen saturation [SpO₂]-based) analogies (SpO₂/Fio₂ and oxygenation saturation index), which have been converted to Pao₂/Fio₂ and oxygenation index using published equations.

TABLE 2. Final Predictive Model for PICU Mortality in the Entire Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology Version 1 Cohort

Variable	OR (95% CI)	Coefficient (95% CI)	Bayesian Information Criterion Increase
Immunocompromised	7.11 (3.96–12.75)	1.961 (1.38–2.55)	37
Pediatric Logistic Organ Dysfunction 2 day 0	1.21 (1.13–1.29)	0.189 (0.12–0.26)	34
Vasopressor-inotrope score day 0 (per 1-point increase)	1.01 (1.01–1.02)	0.014 (0.01–0.02)	13
Pao ₂ /Fio ₂ at 6 hr	0.99 (0.99–0.99)	–0.005 (–0.01 to 0)	12
Fluid balance (mL/kg) day 0	1.01 (1.00–1.01)	0.005 (0–0.01)	11
Constant	–	–2.910 (–3.65 to –2.17)	–

OR = odds ratio.

Increase in Bayesian information criterion (BIC) lists the absolute amount the BIC increases if that particular variable is removed from the final model. Dashes indicate the variable has no interpretable OR.

diagnosis were included. Neither noninvasive ventilation nor presence of unilateral infiltrates were retained as predictors. The model demonstrated good discrimination (AUROC, 0.82; 95% CI, 0.78–0.87), calibration, and fit (Table 3 and Fig. 1A). Ten-fold cross-validation (Table 3) and testing in prespecified subgroups (Supplementary Table 2, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) suggested good internal validity (all AUROC \geq 0.80).

Separate models (Supplementary Table 3, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) were constructed for the cohort excluding those who died of neurologic causes, for those invasively ventilated 6 hours after PARDS diagnosis, and for those invasively ventilated excluding neurologic deaths. Most of the same variables were retained as predictors in all three subgroup models, with the addition of coming from a middle-income country for the cohort excluding neurologic deaths, the substitution of OI (for Pao₂/Fio₂) in the invasively ventilated models, and the removal of day 0 fluid balance in the invasively ventilated models. Retention of OI, rather than Pao₂/Fio₂, resulted in lower BIC in models for intubated subjects. All models (Table 3) demonstrated good discrimination (AUROC \geq 0.84) with retention of this property after cross-validation (cross-validated AUROC \geq 0.83). All models demonstrated good fit (Fig. 1). The models for the intubated cohort had higher AUROC than the severity of illness scores PIM 3, PRISM IV, and PELOD 2, as well as a

published (19) simplified prediction model (all $p < 0.05$ when comparing AUROC; Table 4).

Performance of Models in the CHOP Validation Cohort

Comparison of the PARDIE V1 and the CHOP cohorts are provided in Supplementary Table 4 (Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). By design, the CHOP cohort excluded subjects requiring chronic ventilation at baseline and with unrepaired congenital heart disease, and required subjects to be intubated with bilateral infiltrates. Mortality was similar for both the V1 (17%) and CHOP (18%) cohorts.

Both predictive models (intubated subjects, including and excluding neurologic deaths) had good discrimination (both AUROC \geq 0.85; Table 4) but poor calibration (Supplementary Fig. 1, A and B, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>), with overestimation of mortality when applied without modification to the CHOP cohort. We reasoned that since CHOP represents a tertiary-care center in a developed country, one could reasonably expect differences in both baseline mortality (intercept) as well as the effect of the variables on mortality (slope). Thus, both models were revised in the CHOP dataset using the same variables (Supplementary Table 5, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). The revised models retained good discrimination (both AUROC \geq 0.85), with the revised model for the entire intubated cohort continuing to demonstrate poor fit (Supplementary Fig.

TABLE 3. Predictive Utility, Internal Validation, and Fit of Models

Variable	Whole Cohort	Whole Cohort Excluding Neurologic Deaths		Intubated Excluding Neurologic Deaths
		Whole Cohort	All Intubated	
AUROC ^a	0.82 (0.78–0.87)	0.84 (0.79–0.89)	0.86 (0.81–0.90)	0.90 (0.86–0.94)
AUROC ^a (10-fold cross-validation)	0.81 (0.76–0.86)	0.83 (0.77–0.88)	0.85 (0.80–0.89)	0.84 (0.78–0.90)
Calibration belt p	0.785	0.857	0.838	0.249

AUROC = area under the receiver operating characteristic curve.

^aAll AUROC provided with 95% CIs.

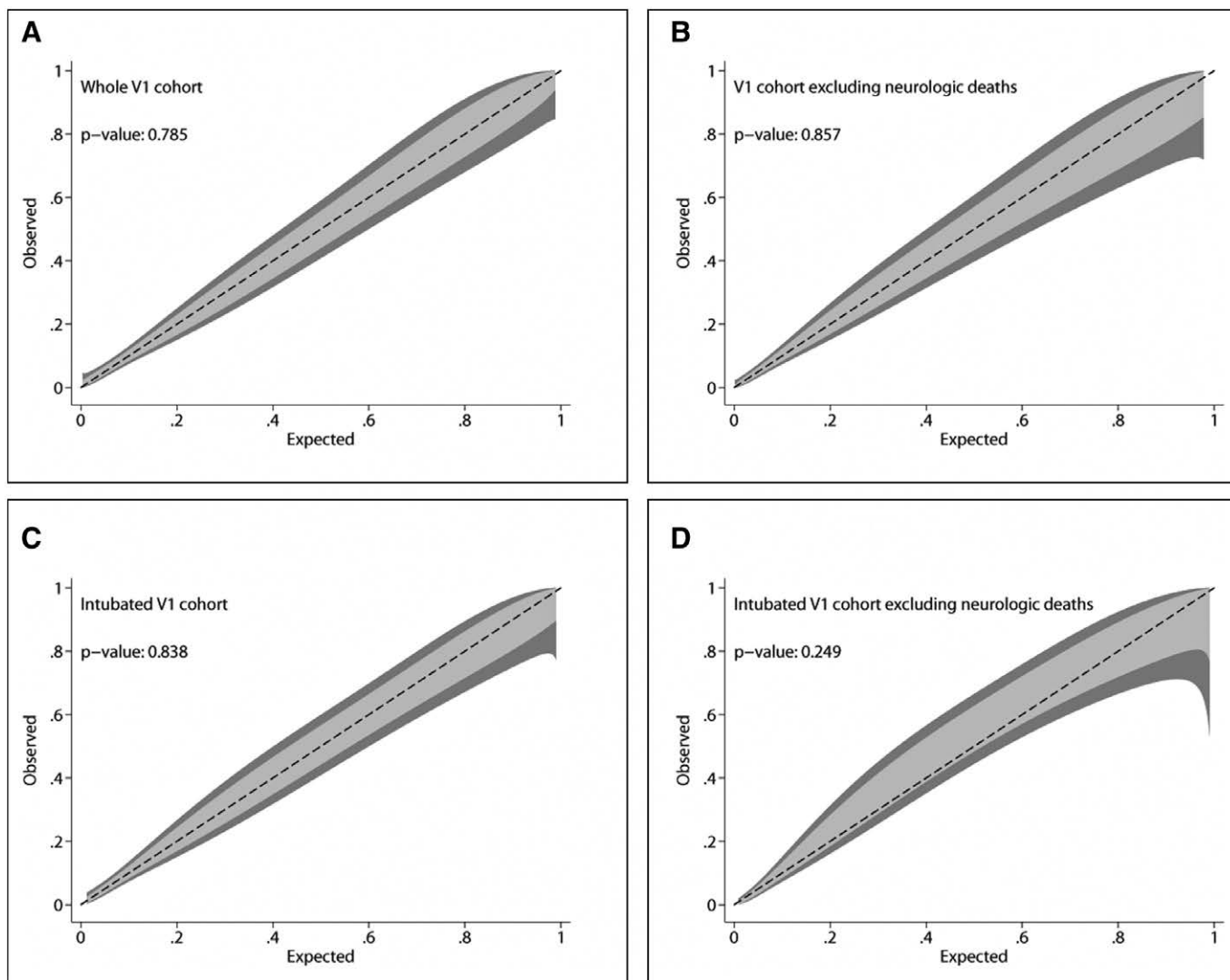


Figure 1. Calibration belts for the separate models predicting PICU mortality in the Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology Version 1 (V1) cohort. The calibration belt examines the relationship between estimated probabilities and observed mortality rates, with associated 80% (*light gray*) and 95% (*dark gray*) CIs. Perfect calibration lies along the center (*dashed*) line. The calibration belt is paired to a statistic that tested deviation from the center line, similar to the Hosmer-Lemeshow test. All four models developed in the V1 cohort demonstrate good calibration (as indicated by $p > 0.05$ in all models).

1C, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). The revised model excluding neurologic deaths demonstrated improved fit (**Supplementary Fig. 1D**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>).

Utility of Mortality Prediction Models for Stratification of VFDs

As VFDs are commonly used as an outcome in PARDS, we assessed whether the four models developed for PICU mortality appropriately stratified VFDs (**Supplementary Fig. 2**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). For models developed in the whole V1 cohort (**Supplementary Fig. 2A**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) and the V1 cohort excluding neurologic deaths (**Supplementary Fig. 2B**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>), we assessed the relationship between quartiles of predicted

mortality and probability of discontinuing total (invasive and noninvasive) ventilation, as not all subjects were intubated. For the models restricted to intubated PARDS (**Supplementary Fig. 2C**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) and intubated PARDS excluding neurologic deaths (**Supplementary Fig. 2D**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>), we assessed the relationship between quartiles of predicted mortality and probability of extubation. For all four models, probability of successful discontinuation of ventilation was appropriately stratified by quartile, with lower probability of discontinuing ventilation (i.e., fewer VFDs) with higher predicted mortality. When we performed a parallel analysis in the CHOP validation cohort of intubated PARDS, both original (**Supplementary Fig. 3, A and B**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) and revised models (**Supplementary Fig. 3, C and D**, Supplemental

TABLE 4. Performance (Area Under the Receiver Operating Characteristic Curve and 95% CI) of the Mortality Prediction Model Relative to Other Severity of Illness Scores and a Published Model for Predicting Pediatric Acute Respiratory Distress Syndrome Mortality

Cohorts	Derivation (Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology Version 1)	Validation (Children's Hospital of Philadelphia)
Intubated cohort ^a		
Model ^b		
Original	0.86 (0.81–0.90)	0.85 (0.81–0.89)
Revised	–	0.85 (0.81 to 0.89)
PIM 3	0.71 (0.65–0.77)	0.70 (0.63–0.78)
PRISM IV	0.77 (0.72–0.82)	0.75 (0.69–0.80)
PELOD 2	0.72 (0.67–0.78)	0.74 (0.68–0.80)
OI at 6 hr + oncology/stem cell transplant	0.76 (0.70–0.82)	0.70 (0.64–0.76)
Intubated excluding neurologic deaths ^a		
Model ^b		
Original	0.90 (0.86–0.94)	0.91 (0.87–0.94)
Revised	–	0.91 (0.87–0.95)
PIM 3	0.68 (0.60–0.75)	0.58 (0.49–0.68)
PRISM IV	0.75 (0.69–0.81)	0.65 (0.57–0.72)
PELOD 2	0.72 (0.65–0.79)	0.71 (0.63–0.79)
OI at 6 hr + oncology/stem cell transplant	0.80 (0.74–0.86)	0.84 (0.79–0.90)

OI = oxygenation index, PELOD = Pediatric Logistic Organ Dysfunction, PIM = Pediatric Index of Mortality, PRISM = Pediatric Risk of Mortality.

^aWe only presented comparisons of the models developed for intubated subjects, as the published simplified pediatric acute respiratory distress syndrome (PARDS) prediction model was restricted to intubated PARDS.

^bWhen comparing area under the receiver operating characteristic curve using Stata "roccomp" (StataCorp, College Station, TX), the predictive model outperformed the severity of illness scores and the simplified prediction model ($p < 0.05$ for all comparisons).

Dashes indicate the the revision only occurred in the Validation cohort.

Digital Content 1, <http://links.lww.com/CCM/F458>) appropriately stratified VFDs.

Models for Ventilator Duration in Survivors

PARDie V1 survivors (**Supplementary Table 6**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) experienced a median of 7.4 days (interquartile range, 4.1–13.0 d) of total (invasive and noninvasive) ventilation, and 6.1 days (interquartile range, 3.1–11.3 d) of invasive ventilation. Factors associated with lower probability of discontinuing total ventilation included higher PELOD 2 and vasopressor score on day of PARDS diagnosis, presence of any comorbidity, and more PICU days before meeting PARDS criteria. Factors associated with lower probability of extubation included vasopressor score on day of PARDS diagnosis, OI at 6 hours after PARDS onset, PICU days before meeting PARDS criteria, and having greater than or equal to 15 PICU beds (**Supplementary Table 7**, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>).

DISCUSSION

PARDS is a heterogeneous syndrome, and severity adjustment tools are important for clinical trial design. To date, there has

not been a readily available and generalizable model for risk stratification in PARDS, with mortality prediction limited to models developed in few centers (13, 19, 20). We developed models for prediction of PICU mortality using variables from the first day of PARDS from a multinational cohort (Table 2; and Appendix 2, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>). The models use readily available clinical data, demonstrate good discrimination and calibration, perform well in subgroups, and outperform published severity of illness scores. Several of the variables in the mortality models are implicated in duration of ventilation, and the mortality prediction model appropriately stratifies VFDs. We externally validated models developed for intubated subjects, thus directly addressing the generalizability concerns of previous models.

A variable selection strategy balancing parsimony and predictive utility protected against over-fitting and increased utility of the models by reducing the number of variables. All of the retained variables have been implicated in PARDS mortality, including immunocompromised status (13, 19) and organ failure (13, 20). Being from a middle-income country was retained as a predictor in the model excluding subjects dying from neurologic causes. This requires further investigation to

identify potential causes of an apparent association between socioeconomic status and mortality. Notably, this is consistent with the lower survival in adults with ARDS from middle-income countries seen in the Large Observational Study to Understand the Global Impact of Severe Acute Respiratory Failure (14). However, it is important to note that only 13% of subjects in PARDIE V1 were from a middle-income country, with complete absence of African representation, limiting any definitive inferences regarding this association.

As in previous studies (16), hypoxemia was responsible for a minority of deaths, whereas MSOF and neurologic etiologies were responsible for a larger fraction. In both adult ARDS (21, 22) and PARDS (8, 16, 20), neurologic etiologies of death are common. Many of these subjects would be excluded from clinical trials due to their neurologic prognosis (22), yet are commonly retained in observational studies, contributing to lower than expected mortality rates in trials. The separate model for intubated patients excluding subjects dying of neurologic causes thus may have greatest utility in clinical trials, as this most closely resembles trial populations, with the caveat that neurologic status may be difficult to assess at PARDS onset.

VFDs are commonly used in trials of adult ARDS and in PARDS. As mortality in pediatrics is lower than adults (12, 23), ventilator duration in survivors is an important contributor to calculating VFDs. For both models predicting ventilator duration (total and invasive ventilation), variables with the highest importance were also present in the models predicting mortality. Overall, this supports the utility of the mortality prediction models for predicting other clinically relevant outcomes, such as VFDs, which we demonstrated in both the derivation (Supplementary Fig. 2, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) and validation (Supplementary Fig. 3, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) cohorts.

Models for intubated subjects were validated in a prospective cohort from CHOP, a large, tertiary-care North American PICU restricted to invasively ventilated subjects with Berlin ARDS, requiring arterial access and bilateral infiltrates. This cohort had higher vasopressor scores, better oxygenation, and more neurologic deaths than the V1 cohort. When comparing the coefficients between the original derivation (Supplementary Table 3, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) and the revised (Supplementary Table 5, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>) models, several coefficients demonstrated downward attenuation. For example, in the revised model for the entire intubated CHOP cohort, the coefficient for PELOD 2 was 27% lower, the coefficient for immunocompromised status was 38% lower, and the coefficient for OI was 34% lower. These results suggest that the impact of these variables on mortality is lower in the CHOP cohort, and that fit remains poor because even this revised model does not capture the best predictors of mortality in this specific PICU, despite identical discrimination (Table 4). Notwithstanding this limitation, the model and variables elected may retain some utility, as both original and revised models appropriately stratify VFDs in the CHOP

cohort (Supplementary Fig. 3, Supplemental Digital Content 1, <http://links.lww.com/CCM/F458>).

The greater percentage of neurologic deaths at CHOP may also contribute to the poor fit, despite revision, of the model developed for all intubated subjects. Notably, the model excluding neurologic deaths demonstrated perfect fit after revision, confirming that cause of death may be impacting the utility of the models. This suggests that, in addition to recalibration or revision, models both including and excluding neurologic deaths need to be tested in a given setting.

We believe that the models offered here are a first step toward improved outcome prediction in PARDS, confirming that a few readily available clinical variables can stratify mortality risk, which will assist with prognostic enrichment for trials (6), and for identifying heterogeneity of treatment effect post hoc (7). We expect further refinement and encourage additional external validation of these models in other PARDS cohorts, particularly in traditionally underrepresented settings.

Strengths of our study include model development in a large, modern, multinational cohort; good internal validation and calibration; and external validation in a separate, large cohort. Our conclusions are based on observations in greater than 1,200 children with PARDS. However, there are limitations to our study. First, variables such as PELOD 2 score, fluid balance, and vasopressor score were collected for the calendar day of PARDS diagnosis, with no adjustment for whether that encompassed 24 hours or a shorter interval. Second, our primary outcome was PICU mortality, rather than hospital mortality, as we reasoned that factors associated with acute PARDS are more likely to affect short-term outcomes, while underlying comorbidities may be more responsible for longer-term outcomes. Third, external validation was only performed in a single-center cohort with different eligibility criteria. Other external validation cohorts which were considered were missing necessary data elements, precluding testing of these models. Accordingly, we believe future PARDS studies should routinely collect these data elements. Fourth, as our focus was on accurate risk stratification for clinical trials, we only tested variables on day 0 of PARDS, which is within the timeframe for recruitment for most studies. The utility of longitudinal models is unknown, but it is possible that these would be an improvement over a static day 0 model. Finally, while an improvement upon existing models, performance of the models for mortality prediction are inadequate to support clinical utility for bedside prognostication at this time.

CONCLUSIONS

We developed and validated predictive models for PICU mortality in PARDS with good performance, which can assist with risk stratification for clinical trials. Development of models excluding neurologic deaths may be of particular utility, as they commonly exclude such subjects. Several variables associated with mortality are also implicated in duration of ventilation, and we confirmed utility of models for stratifying VFDs. All variables represent routinely available clinical data from the first day of PARDS, and should consistently be collected in

future investigations. At this time, model performance is inadequate to support clinical utility at bedside. Studies are warranted to further validate and update these models in other PARDS cohorts, and to test whether operationalized use of the model is feasible in a trial.

Additional members of the Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) Network are: Ira Cheifetz, MD, Ann Thompson, MD, and Scott Watson, MD.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (<http://journals.lww.com/ccmjjournal>).

Supported by Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology funding from University of Southern California Clinical Translational Science Institute; CHU Sainte-Justine, University of Montreal, Canada; Réseau en Santé Respiratoire du Fonds de Recherche Québec-Santé (FRQS); and Children's Hospital Los Angeles, Department of Anesthesiology and Critical Care Medicine.

Dr. Yehya's institution received funding from National Heart, Lung, and Blood Institute and Pfizer. Drs. Yehya, Harhay, Sapru, and Flori received support for article research from the National Institutes of Health (NIH). Dr. Klein received support from University of Southern California Clinical Translational Science Institute; CHU Sainte-Justine, University of Montreal, Canada; Réseau en Santé Respiratoire du Fonds de Recherche Québec-Santé (FRQS); and Children's Hospital Los Angeles, Department of Anesthesiology and Critical Care Medicine. Additional individual funding received from K23-HL136688 (to Dr. Yehya); R00-HL141678 (to Dr. Harhay); Fonds de recherche du Québec Santé (to Dr. Emeriaud); K12-HD047349 (to Dr. Haileselassie); Stanford Maternal Child Health Research Institute Early Career Award (to Dr. Haileselassie); UG3-HL141736 (to Dr. Kneber); UL1-TR000457 (to Dr. Hsing); and Parker B. Francis Fellowship Program (to Dr. Maddux). Dr. Emeriaud's institution received funding from FRQS, and he disclosed that he is the principal investigator of a feasibility study of a new ventilator which is financially supported by Maquet Critical Care. Dr. Flori received funding from Thermo Fisher Scientific (scientific advisory), Aerogen Pharma (scientific advisory), the Society of Critical Care Medicine (travel support), and the Michigan Thoracic Society (Board of Directors). Dr. Maddux's institution received funding from Parker B. Francis Foundation (Fellowship award). Dr. Khemani's institution received funding from Los Angeles Basin Clinical Translational Science Institute through the NIH; received support for article research from Southern California Clinical and Translational Science Institute. The remaining authors have disclosed that they do not have any potential conflicts of interest.

Address requests for reprints to: Nadir Yehya, MD, CHOP, Department of Anesthesiology and Critical Care Medicine, Wood Building 6040A, 34th Street and Civic Center Boulevard, Philadelphia, PA 19104. E-mail: yehyan@email.chop.edu

REFERENCES

- Ashbaugh DG, Bigelow DB, Petty TL, et al: Acute respiratory distress in adults. *Lancet* 1967; 2:319–323
- Bernard GR, Artigas A, Brigham KL, et al: The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994; 149:818–824
- Ranieri VM, Rubenfeld GD, Thompson BT, et al; ARDS Definition Task Force: Acute respiratory distress syndrome: The Berlin definition. *JAMA* 2012; 307:2526–2533
- Khemani RG, Smith LS, Zimmerman JJ, et al; Pediatric Acute Lung Injury Consensus Conference Group: Pediatric acute respiratory distress syndrome: Definition, incidence, and epidemiology: Proceedings from the Pediatric Acute Lung Injury Consensus Conference. *Pediatr Crit Care Med* 2015; 16:S23–S40
- Pediatric Acute Lung Injury Consensus Conference Group: Pediatric acute respiratory distress syndrome: Consensus recommendations from the Pediatric Acute Lung Injury Consensus Conference. *Pediatr Crit Care Med* 2015; 16:428–439
- Prescott HC, Calfee CS, Thompson BT, et al: Toward smarter lumping and smarter splitting: Rethinking strategies for sepsis and acute respiratory distress syndrome clinical trial design. *Am J Respir Crit Care Med* 2016; 194:147–155
- Iwashyna TJ, Burke JF, Sussman JB, et al: Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med* 2015; 192:1045–1051
- Khemani RG, Smith L, Lopez-Fernandez YM, et al; Pediatric Acute Respiratory Distress syndrome Incidence and Epidemiology (PARDIE) Investigators; Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) Network: Paediatric acute respiratory distress syndrome incidence and epidemiology (PARDIE): An international, observational study. *Lancet Respir Med* 2019; 7:115–128
- Khemani RG, Thomas NJ, Venkatachalam V, et al; Pediatric Acute Lung Injury and Sepsis Network Investigators (PALISI): Comparison of SpO₂ to PaO₂ based markers of lung disease severity for children with acute lung injury. *Crit Care Med* 2012; 40:1309–1316
- Parvathaneni K, Belani S, Leung D, et al: Evaluating the performance of the pediatric acute lung injury consensus conference definition of acute respiratory distress syndrome. *Pediatr Crit Care Med* 2017; 18:17–25
- Gaies MG, Gurney JG, Yen AH, et al: Vasoactive-inotropic score as a predictor of morbidity and mortality in infants after cardiopulmonary bypass. *Pediatr Crit Care Med* 2010; 11:234–238
- Yehya N, Servaes S, Thomas NJ: Characterizing degree of lung injury in pediatric acute respiratory distress syndrome. *Crit Care Med* 2015; 43:937–946
- Yehya N, Keim G, Thomas NJ: Subtypes of pediatric acute respiratory distress syndrome have different predictors of mortality. *Intensive Care Med* 2018; 44:1230–1239
- Laffey JG, Madotto F, Bellani G, et al; LUNG SAFE Investigators; ESICM Trials Group: Geo-economic variations in epidemiology, patterns of care, and outcomes in patients with acute respiratory distress syndrome: Insights from the LUNG SAFE prospective cohort study. *Lancet Respir Med* 2017; 5:627–638
- Nattino G, Finazzi S, Bertolini G: A new calibration test and a reappraisal of the calibration belt for the assessment of prediction models based on dichotomous outcomes. *Stat Med* 2014; 33:2390–2407
- Dowell JC, Parvathaneni K, Thomas NJ, et al: Epidemiology of cause of death in pediatric acute respiratory distress syndrome. *Crit Care Med* 2018; 46:1811–1819
- Janssen KJ, Moons KG, Kalkman CJ, et al: Updating methods improved the performance of a clinical prediction model in new patients. *J Clin Epidemiol* 2008; 61:76–86
- Yehya N, Harhay MO, Curley MAQ, et al: Reappraisal of ventilator-free days in critical care research. *Am J Respir Crit Care Med* 2019; 200:828–836
- Spicer AC, Calfee CS, Zinter MS, et al: A simple and robust bedside model for mortality risk in pediatric patients with acute respiratory distress syndrome. *Pediatr Crit Care Med* 2016; 17:907–916
- Flori HR, Glidden DV, Rutherford GW, et al: Pediatric acute lung injury: Prospective evaluation of risk factors associated with mortality. *Am J Respir Crit Care Med* 2005; 171:995–1001
- Stapleton RD, Wang BM, Hudson LD, et al: Causes and timing of death in patients with ARDS. *Chest* 2005; 128:525–532
- Villar J, Martinez D, Mosteiro F, et al; Stratification and Outcome of Acute Respiratory Distress Syndrome (STANDARDS) Network: Is overall mortality the right composite endpoint in clinical trials of acute respiratory distress syndrome? *Crit Care Med* 2018; 46:892–899
- Zimmerman JJ, Akhtar SR, Caldwell E, et al: Incidence and outcomes of pediatric acute lung injury. *Pediatrics* 2009; 124:87–95