Published in final edited form as:

Crit Care Med. 2020 November ; 48(11): e1012–e1019. doi:10.1097/CCM.0000000000004555.

Clinician Accuracy in Identifying and Predicting Organ Dysfunction in Critically Ill Children

Erin F. Carlton, MD, MSc^{1,2}, Jeylan Close, MD^{3,4}, Kelli Paice, MD³, Alyssa Dews, BS⁵, Stephen M. Gorga, MD¹, Julie Sturza, MPH³, Ryan P. Barbaro, MD, MSc^{1,2}, Timothy T. Cornell, MD⁶, Hallie C. Prescott, MD, MSc^{7,8}

¹Department of Pediatrics, Division of Critical Care Medicine, University of Michigan, Ann Arbor, MI

²Susan B. Meister Child Health Evaluation and Research Center, Department of Pediatrics, University of Michigan, Ann Arbor, MI

³Department of Pediatrics, University of Michigan, Ann Arbor, MI

⁴Department of Child and Adolescent Psychiatry and Behavioral Sciences, Children's Hospital of Philadelphia, Philadelphia, PA

⁵College of Literature, Science and the Arts, University of Michigan, Ann Arbor, MI

⁶Department of Pediatrics, Division of Critical Care Medicine, Stanford University, Palo Alto, CA

⁷Department of Internal Medicine, Division of Pulmonary and Critical Care, University of Michigan, Ann Arbor, MI

⁸VA Center for Clinical Management Research, HSR&D Center of Innovation, Ann Arbor, MI

Abstract

Objective: To determine clinician accuracy in the identification and prediction of multiple organ dysfunction syndrome (MODS).

Design: Prospective cohort study

Setting: University of Michigan's C.S. Mott Children's Hospital PICU

Patients: Patients admitted to the PICU with an anticipated PICU length of stay greater than 48 hours.

Intervention: None

Measurement and Results: For each patient, the clinical team (attending, fellow, resident/nurse practitioner) was surveyed regarding existing and anticipated organ dysfunction. The primary outcomes were clinicians' accuracy at identifying MODS and predicting new or

Corresponding author: Erin F. Carlton, MD, Address: 1500 East Medical Center Dr, Ann Arbor, MI 48109, ecarlton@med.umich.edu.

Copyright form disclosure: Drs. Carlton, Barbaro, and Prescott received support for article research from the National Institutes of Health (NIH). Dr. Barbaro disclosed that he is the Extracorporeal Life Support Organization (ELSO) Registry Chair. Dr. Prescott's institution received funding from NIH/National Institute of General Medical Sciences, Agency for Healthcare Research and Quality, and the US Department of Veterans Affairs, and she disclosed government work. The remaining authors have disclosed that they do not have any potential conflicts of interest.

progressive MODS (NP-MODS), compared to the objective assessment of MODS using Proulx criteria. We also measured sensitivity, specificity, negative and positive predictive values, and negative and positive likelihood ratios of clinician assessments. We tested for differences in accuracy by clinician type using Chi square tests. Clinicians rated their confidence in prediction on a 5-point Likert scale.

There were 476 eligible PICU admissions, for whom 1218 surveys were completed. MODS was present in 89 (18.7%) patients at enrollment, and NP-MODS occurred in 39 (8.2%). Clinicians correctly identified MODS with 79.9% accuracy and predicted additional organ dysfunction with 82.6% accuracy. However, the positive and negative likelihood ratios for NP-MODS prediction were 3.0 and 0.7, respectively, indicating a weak relationship between the clinician prediction and development of NP-MODS. The positive predictive value of NP-MODS prediction was just 22.1%.

We found no differences in accuracy by clinician type for either identification of MODS (80.2% vs. 78.2% vs. 81.0%; $p=0.57$) or prediction of NP-MODS (84.8% vs. 82.8% vs. 80.3% $p=0.26$) for attendings, fellows, and residents/nurse practitioners, respectively. There was a weak correlation between the confidence and accuracy of prediction (pairwise correlation coefficient, 0.26; $p<0.001$).

Conclusions: PICU clinicians correctly identified MODS and predicted NP-MODS with 80% accuracy. However, only 8% of patients developed NP-MODS, so accuracy was largely due to true negative predictions. The positive predictive value for NP-MODS prediction was just 22%. Accuracy did not differ by clinician type, but was correlated with self-rated confidence, and was higher for negative predictions.

Keywords

Prediction; Critical Illness; Multiple Organ Dysfunction Syndrome; Pediatrics; Survey

Introduction:

Pediatric Multiple Organ Dysfunction Syndrome (MODS), the dysfunction of two or more concurrent organ systems, occurs in up to half of pediatric ICU (PICU) patients and is associated with both mortality(1-5) and functional decline(1, 6, 7). New or Progressive MODS (NP-MODS) is defined as the development of MODS, or additional organ dysfunctions among patients who already have MODS(8,9). NP-MODS is likewise associated with higher rates of mortality, PICU length of stay, and functional decline(9) .

Pediatric clinical studies increasingly use MODS as a primary outcome due to the low rate of mortality among PICU patients. For example, NP-MODS has been used to assess the impact of the Surviving Sepsis Campaign guidelines(10) and the effect of the age of red blood cell transfusions(11). Both families and clinicians identify organ dysfunction as an important patient-centered outcome(12). Furthermore, clinician identification and prediction of MODS is important to informing treatment decisions in the PICU, such as weighing the risks vs benefit of contrast dye studies, or other diagnostic or therapeutic interventions that might increase risk for new organ dysfunction. Inaccurate assessments and predictions of MODS may therefore lead to increased harm, if risks/benefits of treatments are not weighed

appropriately. To prognosticate accurately, a clinician must both assess the patient's current condition correctly and anticipate their future trajectory.

While prior studies have measured clinical providers' ability to predict mortality(13-17) and functional outcomes(18), clinicians' ability to identify and predict MODS is unclear. To address this gap, we assessed clinicians' ability to identify and predict MODS in a prospective cohort of PICU patients. Furthermore, as prior studies have shown that predictive accuracy for mortality differs by clinician type(13, 16, 19) we tested whether identification and prediction of MODS differ by clinician type. Understanding predictive accuracy and difference in predictive accuracy across clinician types may be helpful for targeting future interventions to improve predictive accuracy.

Methods:

To evaluate clinicians' ability to identify and predict MODS, we performed a prospective observational cohort study of patients admitted to University of Michigan's C.S. Mott Children's Hospital PICU (March 2017 through August 2018). C.S. Mott Children's Hospital is a tertiary hospital with ~1400 PICU admissions annually. PICU patients are cared for by a team of clinicians: one attending, one fellow, and either a nurse practitioner or resident physician.

For consecutive patients admitted to the PICU, a research team member (EFC, JC, or AD) surveyed the PICU clinician team regarding current and anticipated organ dysfunction. The study was approved by University of Michigan IRB. Written informed consent was obtained for each participating clinician and waived for patients.

Clinician Study Population:

Surveys were distributed to each member of the PICU clinical team. In total, 12 attendings, 18 fellows, 92 residents and 9 nurse practitioners completed surveys.

Patient Cohort:

All patients admitted to the PICU with an anticipated PICU stay of 48+ hours were eligible for enrollment, including patients enrolled in hospice or with treatment limitations (e.g. "do not attempt resuscitation" order). We excluded patients with shorter anticipated lengths of stay in order to enrich the cohort with patients at higher risk of developing organ dysfunction. Anticipated duration of PICU stay was assessed by discharge orders, transfer orders, documentation in daily progress notes, or discussion with treating clinicians.

MODS Survey Collection:

Surveys were distributed on 195 days over the 18-month study period, most of which were weekdays. On those days, all eligible patients admitted to the PICU within the last 24 hours were enrolled. Surveys were administered via a web-based platform (Qualtrics, Provo, UT, USA) and typically occurred following morning rounds in the late morning or afternoon, as done in prior studies(16, 20).

Each survey took approximately 2 minutes to complete. Clinicians were asked whether MODS was present, which organs had dysfunction, whether future organ dysfunction would occur in the subsequent week, and which organs would develop dysfunction (eTable 1). Additionally, clinicians rated their degree of certainty regarding predictions of future organ dysfunction on a 5-point Likert scale ranging from 1 “completely uncertain” to 5 “completely certain”.

Objective Assessment of MODS:

The presence of MODS was assessed objectively using Proulx criteria(4, 21) on day of enrollment (baseline MODS), and on days 3, 5, and 7 after enrollment (eTable 2). Proulx criteria has been used in both observational(9, 22) and interventional clinical trials(11) to measure NP-MODS. For each calendar day, the most abnormal value for each parameter was used to determine MODS diagnosis(23). If laboratory values were not collected on a particular day, then the most recent prior value was used to assess MODS criteria.

As in prior studies, new or progressive MODS (NP-MODS) was defined as the development of MODS in the subsequent 7 days (for patients *without* MODS on enrollment) or the development of additional systems of organ dysfunction (for patients *with* MODS on enrollment) (8, 9, 22). We monitored organ function to 7 days since prior studies indicate the majority of organ dysfunction occurs within 7 days of PICU admission(4, 9, 22). Consistent with prior studies, patients who died within 7 days of PICU admission were considered to have NP-MODS(22).

Study Outcomes:

The primary outcomes of this study were overall clinician accuracy of 1) identifying existing MODS, and 2) predicting NP-MODS.

Accuracy was defined as correctly identifying the presence or absence of MODS and was calculated from 2x2 Tables (eTables 3 & 4) as the percentage of true positive and true negatives from the total number of surveys performed. This was determined separately for MODS identification and NP-MODS prediction.

Secondary outcomes were clinician accuracy of identifying each of the 6 individual organ dysfunctions (respiratory, cardiovascular, neurologic, hematologic, hepatic, gastrointestinal); and predicting the future development of these organ dysfunctions.

Power Analysis and Sample Size:

Based on a recent study of PICU patients, we anticipated NP-MODS incidence to be 30%(22). Additionally, we anticipated predictions of NP-MODS to differ by a relative difference of 30% across clinician types, based on a study predicting PICU mortality(16). We thus needed 969 surveys (323 survey per clinician group) to have 80% power to detect a 30% difference in clinician prediction, with an alpha 0.05.

Statistical analysis:

For each outcome, we measured the accuracy, sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, and negative likelihood ratio of clinician assessment.

Similar to prior a study of ICU clinicians (18), we focused primarily on the accuracy, positive likelihood ratio, and negative likelihood ratios as these tests 1) combine sensitivity and specificity and 2) are less influenced by the prevalence of the outcome(24). The likelihood ratio summarizes the probability of the test result in patients who have the disease vs. the probability of the test result in those who do not. A positive likelihood ratio greater than 1 indicates the test is associated with the disease, and a positive likelihood ratio greater than 10 provides strong evidence to rule in disease (18, 25). Similarly, a negative likelihood ratio less than 0.1 provides strong evidence to rule out disease(18, 25).

We tested for differences in accuracy of MODS identification and prediction by clinician type using chi-square tests. Nurse practitioner and resident survey responses were analyzed together since each patient was treated by either a nurse practitioner or by a resident, and their roles are interchangeable. Likewise, to test for differences in the accuracy of confident (4 to 5 on Likert scale) versus non-confident (1 to 3 on Likert scale) predictions, we used chi-square tests. To test whether accuracy of MODS identification was associated with accuracy of NP-MODS prediction, we used McNemar's chi-square tests. We compared certainties of predictions by clinician type using ANOVA. Finally, we used weighted pairwise correlation to measure the association of clinicians' confidence and accuracy of predictions.

Sensitivity Analyses:

We completed a sensitivity analysis in which we accounted for clustering (i.e. multiple surveys per patient) by weighting the surveys such that each PICU admission was weighted equally. 46 patients (9.6%) had more than one PICU admission. We did not adjust for clustering of PICU admissions by patient since each PICU admission represented a distinct opportunity to both identify and predict MODs. In addition, we completed a sensitivity analysis in which we considered only those PICU admissions for which a survey was completed by all 3 clinician groups (attending, fellows and nurse practitioner/resident).

We considered a p-value <0.05 to be significant. Data on objective organ function, Pediatric Risk of Mortality III (PRISM III) score at study enrollment, demographics, diagnoses, PICU treatments, length of stay, and mortality were collected from the electronic health record. Data were collected and managed using REDCap electronic data capture tools hosted at the University of Michigan(26, 27). Analyses were conducted in StataMP software Version 15 (StataCorp, College Station, TX) or 9.4 (SAS Institute Inc., Cary, NC).

Results:

During the 195 study days, there were 476 eligible PICU admissions, for whom 1218 surveys were completed (median 3 surveys per PICU admission (IQR 2-3)). Of the 1218 total surveys, 348 (28.6%) were completed by attendings, 454 (37.3%) by fellows, and 416

(34.1%) by nurse practitioners/ residents. 226 (47.5%) PICU admissions had surveys completed by each clinician type.

The patients were 52% male with a median age of 4 years (IQR 1-12, range 0-21). The median PICU length of stay was 3 days (IQR 2-8), while overall hospital length of stay was 8 days (IQR 4-16). Median PRISM score at enrollment was 2 (IQR 0-5). 39.7% of patients were mechanically ventilated, 13.2% received vasopressors, and 17 (3.6%) died during the PICU admission (Table 1).

Presence of Organ Dysfunction and MODS by Proulx criteria:

MODS was present in 89 (18.7%) patients on enrollment. By day 7, 62 patients (13.0%) developed a new organ dysfunction. Of these, 17 (3.6% of all patients) met criteria for new MODS, and 22 (4.6%) met criteria for progressive MODS. In total, 106 (22.2%) had MODS during the first 7 days of PICU admission, and 39 (8.2%) had NP-MODS (Figure 1).

The most common organ dysfunctions at enrollment were respiratory (39.5% of all patients), cardiovascular (11.5%), and neurologic (10.9%) (Figure 2). The most common subsequent organ dysfunctions, developing in the week after enrollment, were respiratory (5.5% of all patients), cardiovascular (3.4%), and renal (2.9%) dysfunction.

Clinician Identification of MODS

There were 231 surveys completed for patients with MODS at enrollment, of which 124 (53.7%) correctly identified MODS. Among the 984 surveys in patients without baseline MODS, clinicians correctly identified the absence of MODS in 847 (86.1%). In total, clinicians identified the presence or absence of MODS with 79.9% accuracy (95% CI 77.6%, 82.1%) (Table 2). In the sensitivity analysis using weighted data, clinicians identified MODS with 80.2% accuracy (95% CI 76.6%, 83.9%) (eTable 5).

Clinicians incorrectly identified MODS in 11.2% of surveys (95% CI 9.5%, 13.1%), indicating MODS was present when it was not. Conversely, clinicians failed to identify MODS when it was actually present in 8.8% of surveys (95% CI 7.2%, 10.5%). Overall, the positive predictive value for clinician identification of MODS was 47.5%, with a false negative rate (1-sensitivity) of 46.3%.

Clinician Prediction of NP-MODS

Of 105 surveys completed for patients who developed NP-MODS, clinicians correctly predicted NP-MODS in 42 (40.0%). Among 1111 surveys completed for patients without NP-MODS, clinicians correctly predicted the lack of NP-MODS in 963 (86.7%). Overall, clinicians correctly predicted the presence or absence of NP-MODS with 82.6% accuracy (95% CI 80.1%, 84.7%). In the sensitivity analysis using weighted data, clinicians predicted NP-MODS with 83.6% accuracy (95% CI 79.9%, 86.8%) (eTable 5).

Clinicians incorrectly predicted the development of MODS in 12.1% of surveys (95% CI 10.4%, 14.1%), indicating it would develop when it did not. Conversely, in 5.2% of surveys (95% CI 3.9, 6.6), clinicians predicted no NP-MODS but NP-MODS occurred. The positive likelihood ratio for prediction of NP-MODS was 3.0, while the negative likelihood ratio was

0.7 (Table 2), demonstrating a weak relationship between clinician prediction and the development of NP-MODS. The positive predictive value for NP-MODS prediction was 22.1%, with a false-negative rate of 60%.

Prediction of NP-MODS was accurate in 85.5% of surveys for which identification of MODS was correct. However, of the 247 surveys that did not correctly identify MODS on enrollment, 70.4% (n=174) correctly predicted if a patient would or would not develop NP-MODS. Overall, there was not a significant association between a correct identification and subsequent prediction of MODS (p=0.06).

Identification and Prediction by Clinician Type

MODS was identified with 80.2% accuracy by attendings, 78.2% by fellows, and 81.0% by nurse practitioners/residents (p=0.57). NP-MODS was predicted with 84.8% accuracy by attendings, 82.8% accuracy by fellows, and 80.3% accuracy by nurse practitioners/residents (p=0.26).

Sensitivity Analysis

Of the 476 PICU admissions in our cohort, 226 (47.5%) had surveys completed by all 3 clinician groups (attendings, fellows, nurse practitioners/residents). Similar to the complete cohort, 15.9% (n=36) had MODS on enrollment and 8.4% (n=19) developed NP-MODS. Clinicians correctly identified MODS with 80.4% accuracy and predicted NP-MODS with 82.6% accuracy (Online Supplement eTable 6), similar to the primary analysis. There was likewise no difference in accuracy of identification or prediction of MODS between clinician types (Online Supplement eTable 7).

Clinician Identification and Prediction of Individual Organ Dysfunctions

Clinicians' accuracy for identifying individual organ dysfunctions was greater than 90% for gastrointestinal, hepatic, renal and hematological dysfunctions, 85.1% for neurologic dysfunction, and 65.0% for respiratory dysfunction (eTable 3). Clinicians' accuracy for predicting future organ dysfunction was greater than 90% for gastrointestinal, hepatic, hematologic, neurologic, and cardiovascular dysfunctions, 85.1% for respiratory dysfunction, and 83.7% for renal dysfunction (eTable 4).

Confidence of Predictions, and Association with Accuracy

Confidence in predicting the development of NP-MODS was highest among attendings (mean 3.9 on 5-point Likert scale, median 4), followed by residents/nurse practitioners (mean 3.7, median 4) then fellows (mean 3.5, median 4) (p<0.001) (Figure 3). Overall, 60.8% of predictions were confident (4-5 on 5-point Likert Scale). However, among 189 surveys predicting the development of NP-MODS, only 59 (31.2%) were confident.

There was a weak correlation between the confidence and accuracy of prediction of NP-MODS (weighted pairwise correlation coefficient, 0.26; p<0.001). Clinician predictions of NP-MODS were 89.4% accurate when rated as confident, compared to 70.4% accurate when not rated as confident, p<0.01.

Discussion

In this prospective cohort study of nearly 500 PICU patients, we found clinicians correctly identified MODS and predicted the development of NP-MODS with approximately 80% accuracy. However, as the rate of MODS and NP-MODS were low (20% and 7%, respectively), the sensitivity and positive predictive value of identification and prediction were low. Accuracy of assessments was similar across clinician types, but prediction of NP-MODS differed significantly from objective assessment using Proulx criteria.

The number of false positive and false negative responses were similar when identifying MODS. In contrast, clinicians were more pessimistic when asked to predict the development of NP-MODS. In 12% of surveys, clinicians predicted NP-MODS that did not occur (ie, a false positive prediction). By contrast, in 5% of surveys, clinicians failed to predict NP-MODS that occurred (ie, false negative prediction). These findings are consistent with prior studies showing that false positive predictions of pediatric (16) and adult (13, 17, 28) mortality are more common than false negative predictions. Clinicians may have overestimated the development of NP-MODs for several reasons. First, clinicians may err on the side of predicting worse outcomes, in order to prepare families for this possibility. Patients and families are likely to respond better to having the child fare better than they were counseled, compared to faring worse. Second, this overestimation could be due to clinicians recognizing and predicting worsening organ dysfunction by the spirit rather than the letter of the law. For example, a clinician may identify or predict organ dysfunction based on a negative clinical trajectory, even if the patient did not reach the Proulx criteria threshold for organ dysfunction.

Another important finding of this study is that MODS and NP-MODS were relatively uncommon among this cohort of all-cause admissions to a tertiary PICU. Of the 478 patients in our study, around 20% had baseline MODS, with an additional 7% developing NP-MODS within one week. By contrast, prior studies reported NP-MODS in 10-30% of PICU patients (9, 11, 22, 29). We suspect these differences are explained by study inclusion/exclusion criteria. The prior studies focused on children with severe sepsis or acute respiratory failure (9, 11, 22, 29), while our study included all-cause PICU admissions.

It should be noted that this low prevalence of MODS and NP-MODS contributed to the overall accuracy of clinician identification and prediction. As accuracy is the percentage of true positive and true negatives from the total number of surveys performed, the low prevalence of disease means the true negatives make up a larger portion of the numerator. Indeed, while the accuracy was approximately 80% for MODS identification and prediction, the positive likelihood ratios were less than 10 demonstrating a weak association between the clinician identification or prediction and the probability of MODS/NP-MODS. The impact of low prevalence is especially highlighted in the large discrepancy between positive predictive (PPV) and negative predictive values (NPV)—MODS 47.5% PPV vs. 88.8% NPV and NP-MODS 22.1% PPV vs 93.9% NPV.

Accuracy of predictions differed by organ system. Clinicians correctly identified and predicted gastrointestinal, hepatic, and hematologic failure with more than 90% accuracy.

Clinicians were least accurate in their identification of respiratory failure (65% accuracy) and prediction of future renal failure (84% accuracy). The lower accuracy for these organ dysfunctions may be in part due to their higher frequency. Indeed, respiratory failure treated with invasive or non-invasive ventilation occurs in half of PICU patients worldwide(30), while acute kidney injury occurs in more than 25% (31). Additionally, both respiratory and renal dysfunction may be identified in clinical practice using a variety of criteria besides Proulx (e.g. Pediatric Acute Respiratory Distress Syndrome(32) and Kidney Disease: Improving Global Outcomes Acute Kidney Injury definitions(33)). Thus, clinicians' identification of respiratory and renal dysfunction may be more likely to differ from the Proulx criteria given the multiple definitions used in clinical practice.

Our study differed from prior studies which showed predictive accuracy differs across clinician types(16, 19). For example, a prior study of 600 PICU patients found that attending physicians predicted mortality with greater accuracy than fellows, residents and nurses(16). Similarly, in a study of terminally ill adults, physician experience was associated with prediction accuracy(19). We distributed surveys following morning rounds, so clinical discussion on rounds may have allowed the team to reach similar conclusions regarding MODS diagnosis and risk of NP-MODS. Additionally, as less than 20% of the cohort had MODS or NP-MODS, it is possible that the sample size was too small to detect a difference.

A clinician's self-rating of confidence is correlated with accuracy of prediction (16, 18). In our cohort, clinicians were confident in over 60% of the surveys. However, among surveys that predicted the development of NP-MODS, clinicians were confident only a third of the time. This is consistent with prior studies showing that clinicians tend to be more confident when predicting patients to be at low risk of an event (33). Additionally, accuracy of prediction was nearly 20% higher among surveys with high clinician confidence. Overall, clinicians were more confident when predicting patients would not develop additional organ dysfunction compared to predicting patients would acquire additional organ dysfunction. Thus, tools designed to identify patients who may develop organ failure would be of greater benefit compared to those that rule out future organ failure.

The better clinicians can anticipate the development of new organ failure, the better they can weigh the risks/benefits of diagnostic or therapeutic interventions and decide the optimal management plan for an individual patient. Previous studies comparing clinician intuition to prediction models for mortality have shown clinicians perform better (34, 35), or equally as well(28, 36). However, clinician predictions may still be improved when combined with a prediction model(36). In our study, clinicians correctly predicted the development of additional organ failure only 40% of the time—suggesting an opportunity to improve discriminative accuracy with predictive tools. However, for such tools to be useful, they must be easily applied at the bedside.

This study has several limitations. First, it was performed at a single institution, so the findings may not be generalizable to other settings. However, we collected more than 1200 surveys of 131 clinicians over an 18-month period. Second, it is possible that we did not capture all patients who developed NP-MODS as we excluded patients with an anticipated PICU length of stay less than 48 hours. However, given the low rate of NP-MODS among

children who did remain in the PICU, it is unlikely that many NP-MODS events were excluded by this sampling strategy. Third, surveys were typically collected on weekdays, thus potentially limiting the cohort size. Fourth, because the rate of NP-MODS was lower than anticipated based on prior literature, our power to detect differences in predictive accuracy by clinician types was limited. Finally, we measured objective organ dysfunction by Proulx on days 3, 5, and 7 after enrollment, so may have missed transient or late organ dysfunction.

Conclusion

PICU clinicians correctly identified MODS and predicted NP-MODS with 80% accuracy. However, the sensitivity and positive predictive values were low, suggesting further opportunity to improve predictive ability. Accuracy of identification and predictions did not differ by clinician type, but was correlated with self-rated confidence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding:

This material is the result of work supported with resources and use of facilities at the Ann Arbor VA Medical Center. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the U.S. government. This work was supported by grants K08 GM115859 [HCP] and K12 HL138039 [RBP] from the National Institutes of Health.

References

1. Typpo KV, Petersen NJ, Hallman DM, et al.: Day 1 multiple organ dysfunction syndrome is associated with poor functional outcome and mortality in the pediatric intensive care unit. *Pediatr Crit Care Med* 2009; 10:562–570 [PubMed: 19741445]
2. Leteurtre S, Duhamel A, Salleron J, et al.: PELOD-2: an update of the PEdiatric logistic organ dysfunction score. *Crit Care Med* 2013; 41:1761–1773 [PubMed: 23685639]
3. Kutko MC, Calarco MP, Flaherty MB, et al.: Mortality rates in pediatric septic shock with and without multiple organ system failure. *Pediatr Crit Care Med* 2003; 4:333–337 [PubMed: 12831416]
4. Proulx F, Gauthier M, Nadeau D, et al.: Timing and predictors of death in pediatric patients with multiple organ system failure. *Crit Care Med* 1994; 22:1025–1031 [PubMed: 8205810]
5. Leclerc F, Leteurtre S, Duhamel A, et al.: Cumulative Influence of Organ Dysfunctions and Septic State on Mortality of Critically Ill Children. *Am J Resp Crit Care* 2005; 171:348–353
6. Typpo K, Watson RS, Bennett TD, et al.: Outcomes of Day 1 Multiple Organ Dysfunction Syndrome in the PICU. *Pediatr Crit Care Med* 2019; 20:914–922 [PubMed: 31589198]
7. Farris RWD, Weiss NS, Zimmerman JJ: Functional Outcomes in Pediatric Severe Sepsis. *Pediatr Crit Care Med* 2013; 14:835–842 [PubMed: 24108117]
8. Villeneuve A, Joyal J-S, Proulx F, et al.: Multiple organ dysfunction syndrome in critically ill children: clinical value of two lists of diagnostic criteria. *Ann Intensive Care* 2016; 1–7 [PubMed: 26728593]
9. Lin JC, Spinella PC, Fitzgerald JC, et al.: New or Progressive Multiple Organ Dysfunction Syndrome in Pediatric Severe Sepsis: A Sepsis Phenotype With Higher Morbidity and Mortality. *Pediatr Crit Care Med* 2017; 18:8–16 [PubMed: 28060151]

10. Workman JK, Ames SG, Reeder RW, et al.: Treatment of Pediatric Septic Shock With the Surviving Sepsis Campaign Guidelines and PICU Patient Outcomes*. *Pediatr Crit Care Med* 2016; 17:e451–e458 [PubMed: 27500722]
11. Spinella PC, Tucci M, Fergusson DA, et al.: Effect of Fresh vs Standard-issue Red Blood Cell Transfusions on Multiple Organ Dysfunction Syndrome in Critically Ill Pediatric Patients. *JAMA* 2019; 322:2179–2190 [PubMed: 31821429]
12. Merritt C, Menon K, Agus MSD, et al.: Beyond Survival. *Pediatr Crit Care Med* 2018; 19:e105–e111 [PubMed: 29394234]
13. Rocker G, Cook D, Sjøkvist P, et al.: Clinician predictions of intensive care unit mortality*. *Crit Care Med* 2004; 32:1149–1154 [PubMed: 15190965]
14. McClish DK, Powell SH: How Well Can Physicians Estimate Mortality in a Medical Intensive Care Unit? *Med Decis Making* 1989; 9:125–132 [PubMed: 2747449]
15. Nadler I, Sanderson PM: Using Brunswik's Probabilistic Functionalism to Test How Clinicians Make Judgments in Simulated Neonatal Resuscitation Scenarios. *Pediatrics* 1994; 93:945–950 [PubMed: 8190582]
16. Marcin JP, Pollack MM, Patel KM, et al.: Prognostication and Certainty in the Pediatric Intensive Care Unit. *Pediatrics* 1999; 104:868–873 [PubMed: 10506227]
17. Brabrand M, Hallas J, Knudsen T: Nurses and Physicians in a Medical Admission Unit Can Accurately Predict Mortality of Acutely Admitted Patients: A Prospective Cohort Study. *PLOS ONE* 2014; 9:e101739–7 [PubMed: 25019354]
18. Detsky ME, Harhay MO, Bayard DF, et al.: Discriminative Accuracy of Physician and Nurse Predictions for Survival and Functional Outcomes 6 Months After an ICU Admission. *JAMA* 2017; 317:2187–2195 [PubMed: 28528347]
19. Christakis NA, Lamont EB, Smith JL, et al.: Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study Commentary: Why do doctors overestimate? Commentary: Prognoses should be based on proved indices not intuition. *BMJ* 2000; 320:469–473 [PubMed: 10678857]
20. Chopra V, Govindan S, Kuhn L, et al.: Do Clinicians Know Which of Their Patients Have Central Venous Catheters?: A Multicenter Observational Study. *Ann Intern Med* 2014; 161:562 [PubMed: 25329204]
21. Proulx F, Fayon M, Farrell CA, et al.: Epidemiology of Sepsis and Multiple Organ Dysfunction Syndrome in Children. *Chest* 1996; 109:1033–1037 [PubMed: 8635327]
22. Weiss SL, Fitzgerald JC, Pappachan J, et al.: Global epidemiology of pediatric severe sepsis: the sepsis prevalence, outcomes, and therapies study. *Am J Resp Crit Care* 2015; 191:1147–57
23. Bestati N, Leteurtre S, Duhamel A, et al.: Differences in organ dysfunctions between neonates and older children: a prospective, observational, multicenter study. *Crit Care* 2010; 14:R202 [PubMed: 21062434]
24. Jaeschke R, Guyatt GH, Sackett DL, et al.: Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271:703–707 [PubMed: 8309035]
25. Deeks JJ, Altman DG: Diagnostic tests 4: likelihood ratios. *BMJ* 2004; 329:168 [PubMed: 15258077]
26. Harris PA, Taylor R, Thielke R, et al.: Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; 42:377–381 [PubMed: 18929686]
27. Harris PA, Taylor R, Minor BL, et al.: The REDCap Consortium: Building an International Community of Software Platform Partners. *J Biomed Inform* 2019; 95:103208 [PubMed: 31078660]
28. Litton E, Ho KM, Webb SA: Comparison of physician prediction with 2 prognostic scoring systems in predicting 2-year mortality after intensive care admission: A linked-data cohort study. *J Crit Care* 2012; 27:423.e9–423.e15

29. Weiss SL, Asaro LA, Flori HR, et al.: Multiple Organ Dysfunction in Children Mechanically Ventilated for Acute Respiratory Failure. *Pediatr Crit Care Med* 2017; 18:319–329 [PubMed: 28212163]
30. Khemani RG, Smith L, Lopez-Fernandez YM, et al.: Paediatric acute respiratory distress syndrome incidence and epidemiology (PARDIE): an international, observational study. *Lancet Respir Med* 2019; 7:115–128 [PubMed: 30361119]
31. Kaddourah A, Basu RK, Bagshaw SM, et al.: Epidemiology of Acute Kidney Injury in Critically Ill Children and Young Adults. *New Engl J Med* 2017; 376:11–20 [PubMed: 27959707]
32. Group PALICC: Pediatric acute respiratory distress syndrome: consensus recommendations from the Pediatric Acute Lung Injury Consensus Conference. *Pediatr Crit Care Med* 2015; 16:428–39 [PubMed: 25647235]
33. Kellum JA, Lameire N: Diagnosis, evaluation, and management of acute kidney injury: a KDIGO summary (Part 1). *Crit Care* 2013; 17:204 [PubMed: 23394211]
34. Sinuff T, Adhikari NKJ, Cook DJ, et al.: Mortality predictions in the intensive care unit: Comparing physicians with scoring systems*. *Crit Care Med* 2006; 34:878–885 [PubMed: 16505667]
35. Farion KJ, Wilk S, Michalowski W, et al.: Comparing predictions made by a prediction model, clinical score, and physicians. *Appl Clin Inform* 2013; 04:376–391
36. Marcin JP, Pollack MM, Patel KM, et al.: Combining physician's subjective and physiology-based objective mortality risk predictions. *Crit Care Med* 2000; 28:2984–2991 [PubMed: 10966283]

Figure 1:
Flowchart of Enrollment and New or Progressive Multiple Organ Dysfunction Syndrome

Figure 2: Clinician Identification and Prediction of Organ Dysfunction on Enrollment and New Organ Dysfunctions

These stacked bar graphs represent the percentage of survey responses for identification and prediction of Multiple Organ Dysfunction Syndrome (MODS) and New or Progressive MODS (NP-MODS) at baseline and by day 7 of pediatric intensive care unit admission. Each individual organ system is also presented. The yellow bars represent those patients with organ dysfunction, while blue bars represent those without.

Figure 3: Self-Reported Confidence in Prediction of Organ Dysfunction by Clinicians
For certainty ratings, 1 indicates not very certain, while 5 indicates very certain that the patient will or will not develop organ dysfunction within the next 7 days.

Table 1.**Patient Characteristics**

Characteristics	Patients, No. (%) n=476
Age, median (IQR), year	4 (1-12)
Female	230 (48.4%)
Race	
White	359 (75.4%)
African American	80 (16.8%)
Hispanic	27 (5.7%)
Infection	249 (52.3%)
Bacterial	119 (25.0%)
Viral	178 (37.2%)
Fungal	10 (2.1%)
Primary Diagnosis	
Respiratory Failure/Distress	192 (40.3%)
Sepsis/Septic Shock *	32 (6.7%)
Seizure or Altered Mental Status	26 (5.4%)
Mechanical Ventilation	189 (39.7%)
Vasopressors	63 (13.2%)
Continuous Renal Replacement Therapy	14 (2.9%)
Extracorporeal Membrane Oxygenation	9 (1.9%)
PRISM Score at study enrollment, median (IQR)	2 (0-5)
PICU Length of Stay, median (IQR), days	3 (2-8)
Hospital Length of Stay median (IQR), days	8 (4-16)
PICU Mortality	17 (3.6%)
Hospital Mortality	21 (4.4%)

* 18 additional patients with documented secondary diagnosis of sepsis/shock, with total of 10.5% of patients with primary or secondary diagnosis of sepsis/septic shock

Table 2.

Performance of Clinician Identification and Prediction, overall and by clinician type

	Accuracy	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Likelihood Ratio	
						Positive	Negative
Identification of Current MODS[*]:							
<i>Overall</i>	79.9%	53.7%	86.1%	47.5%	88.8%	3.9 (3.2,4.7)	0.5 (0.5,0.6)
<i>Attending Physicians</i>	80.2%	61.9%	84.2%	46.4%	90.9%	3.9 (2.8,5.5)	0.5 (0.3,0.6)
<i>Fellow Physicians</i>	78.2%	51.1%	85.4%	46.5%	87.5%	3.5 (2.5,4.8)	0.6 (0.5,0.7)
<i>Nurse Practitioners or Resident Physicians</i>	81.0%	50.0%	88.4%	50.0%	88.4%	4.3 (3.0,6.3)	0.6 (0.5,0.7)
Prediction of NP-MODS^{**}:							
<i>Overall</i>	82.6%	40.0%	86.7%	22.1%	93.9%	3.0 (2.3,4.0)	0.7 (0.6,0.8)
<i>Attending Physicians</i>	84.8%	37.9%	89.3%	24.4%	94.0%	3.6 (2.0,6.2)	0.7 (0.5,0.9)
<i>Fellow Physicians</i>	82.8%	36.8%	87.0%	20.6%	93.8%	2.8 (1.8,4.6)	0.7 (0.6,0.9)
<i>Nurse Practitioners or Resident Physicians</i>	80.3%	44.7%	84.1%	22.1%	93.8%	2.8 (1.8,4.3)	0.7 (0.5,0.9)

* Multiple Organ Dysfunction Syndrome occurred in 89 (18.7%) PICU admissions

** New or Progressive Multiple Organ Dysfunction Syndrome occurred in 39 (8.2%) PICU admissions