Pediatric Index of Cardiac Surgical Intensive Care Mortality Risk Score for Pediatric Cardiac Critical Care*

Howard E. Jeffries, MD, MBA[1]; Gerardo Soto-Campos, PhD[2]; Aaron Katch, MS[2]; Christine Gall, DrPH, MS, RN[2]; Tom B. Rice, MD[3]; Randall Wetzel, MD[4]

Objective: Comparison of clinical outcomes is imperative in the evaluation of healthcare quality. Risk adjustment for children undergoing cardiac surgery poses unique challenges, due to its distinct nature. We developed a risk-adjustment tool specifically focused on critical care mortality for the pediatric cardiac surgical population: the Pediatric Index of Cardiac Surgical Intensive care Mortality score.

Design: Retrospective analysis of prospectively collected pediatric critical care data.

Setting: Pediatric critical care units in the United States.

Patients: Pediatric cardiac intensive care surgical patients.

Interventions: Prospectively collected data from consecutive patients admitted to ICUs were obtained from The Virtual PICU System (VPS, LLC, Los Angeles, CA), a national pediatric critical care database. Thirty-two candidate physiologic, demographic, and diagnostic variables were analyzed for inclusion in the development of the Pediatric Index of Cardiac Surgical Intensive care Mortality model. Multivariate logistic regression with stepwise selection was used to develop the model.

Measurements and Main Results: A total of 16,574 cardiac surgical patients from the 55 PICUs across the United States were included in the analysis. Thirteen variables remained in the final model, including the validated Society of Thoracic Surgeons-European Association of Cardio-Thoracic Surgery Congenital Heart Surgery Mortality (STAT) score and admission time with respect to cardiac surgery, which identifies whether the patient underwent the index surgical procedure before or after admission to the ICU. Pediatric Index of Cardiac Surgical Intensive Care Mortality (PICSIM) performance was compared with the performance of Pediatric Risk of Mortality-3 and Pediatric Index of Mortality-2 risk of mortality scores, as well as the STAT score and STAT categories by calculating the area under the curve of the receiver operating characteristic from a validation dataset: PICSIM (area under the curve = 0.87) performed better than Pediatric Index of Mortality-2 (area under the curve = 0.81), Pediatric Risk of Mortality-3 (area under the curve = 0.82), STAT score (area under the curve = 0.77), STAT category (area under the curve = 0.75), and Risk Adjustment for Congenital Heart Surgery-1 (area under the curve = 0.74).

Conclusions: This newly developed mortality score, PICSIM, consisting of 13 risk variables encompassing physiology, cardiovascular condition, and time of admission to the ICU showed better discrimination than Pediatric Index of Mortality-2, Pediatric Risk of Mortality-3, and STAT score and category for mortality in a multisite cohort of pediatric cardiac surgical patients. The introduction of the variable "admission time with respect to cardiac surgery" allowed prediction of mortality when patients are admitted to the ICU either before or after the index surgical procedure. (*Pediatr Crit Care Med* 2015; 16:846–852)

Key Words: database analytics; mortality prediction; pediatric cardiac intensive care; pediatric cardiac surgery; predictive modeling; quality improvement

***See also p. 885.**

[1]Department of Pediatrics, Seattle Children's Hospital, University of Washington School of Medicine, Seattle, WA.

[2]Virtual PICU System, LLC, Los Angeles, CA.

[3]Department of Pediatrics, Children's Hospital of Wisconsin, Medical College of Wisconsin, Milwaukee, WI.

[4]Department of Anesthesiology & Critical Care Medicine, Children's Hospital Los Angeles, University of Southern California School of Medicine, Los Angeles, CA.

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/pccmjournal).

Dr. Jeffries disclosed that he is a clinical consultant to the VPS, LLC and receives nominal compensation for his allocated time. Dr. Gall is an employee of VPS and in the course of his daily activities participated in this research. He received no specific reimbursement for this work other than my standard paycheck. Dr. Rice served as a board member for VPS, LLC (uncompensated) and is employed by the Medical College of Wisconsin. His institution consulted for NACHRI/CHA (QTN/CABSI Collaborative, National Faculty member), received grant support from CHIPRA, and received support for travel for the NACHRI CABSI Collaborative (national faculty and travel to workshops). Dr. Wetzel is the uncompensated CEO of VPS, LLC (mentioned in the paper). The remaining authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: howard.jeffries@seattlechildrens.org

DOI: 10.1097/PCC.0000000000000489

Comparing healthcare outcomes among populations is an important element in the evaluation of healthcare quality and benchmarking. Increasingly, providers, consumers, administrators, policy makers, and payers are demanding evidence that healthcare services are delivered effectively and efficiently and in accordance with current standards of practice. Importantly, the comparison of outcomes among centers requires adjusting for some measure of patients' or population's severity of illness, such as the population's risk of mortality to avoid inaccurate conclusions regarding quality of care and center performance. In addition, studying populations for comparative effectiveness research or to improve quality by comparing therapeutic approaches requires the assurance that the populations being studied are comparable with regard to severity of illness. Reliable comparative analysis of outcomes depends on a clear understanding of the factors that influence the risk of mortality for populations of critically ill children so that proper risk adjustment may occur (1).

In pediatric critical care, several physiologic-based scoring systems have been developed for risk adjustment within clinical datasets and among ICUs (2–4). These tools enable the establishment of baseline severity-of-illness measurement that provides an estimated mortality for the critically ill population being evaluated. The predicted mortality may then be compared with the actual mortality to calculate a standardized mortality ratio (SMR) to account for the variation in severity of illness (5). There is evidence that the performance of generic risk-adjustment tools suffers when specific patient populations (cardiac, postoperative, leukemia, etc.) are not similar to the heterogeneous population in which they were developed (6). As such, pioneers in the field of risk adjustment suggest that diagnosis-specific scores may be advantageous when comparing the outcomes of multiple ICUs (5–9).

Risk adjustment for pediatric populations undergoing cardiac surgery poses unique challenges. In addition to the wide variability in surgical case complexity, variations in anatomy may dramatically alter the degree of difficulty for a given surgical procedure. Comorbidities, which are frequently encountered in patients with congenital heart disease, may often increase the risk of mortality, independent of the surgical procedure being performed. Existing cardiac severity-of-illness scores, such as Risk Adjustment for Congenital Heart Surgery (RACHS), Society of Thoracic Surgeons-European Association of Cardio-Thoracic Surgery Congenital Heart Surgery Mortality (STAT) scores, and Aristotle scores, assess overall cardiac surgical mortality across the entire care process but do not assess severity of illness at admission to the ICU. Due to the distinct nature of congenital heart surgery intensive care, and with the advent of dedicated cardiac units, scores developed for ICU cardiac surgical populations could be expected to perform better than nonspecific scores for comparing populations outcomes and unit performance; we sought to develop a risk-adjustment tool specifically for the pediatric cardiac surgical intensive care population, which we have termed as the Pediatric Index of Cardiac Surgical Intensive Care Mortality (PICSIM) score.

MATERIALS AND METHODS

Study Sites

Data were obtained from the Virtual PICU System (VPS) database (VPS, LLC, Los Angeles, CA; http://www.myvps.org), a national pediatric critical care database. The study was deemed to be exempt from approval by the Seattle Children's Institutional Review Board. VPS data were provided by the VPS, LLC. No endorsement or editorial restriction of the interpretation of these data or opinions of the authors has been implied or stated. Data entered into VPS are entered by trained data collectors using standardized data definitions and routinely assessed for interrater reliability, which was greater than 93% concordance for all data collection periods included in this study. After data submission, all data were reviewed and validated prior to inclusion in the dataset used for quality measures and for research studies. This study focused on children with congenital heart defects who had cardiac surgery and were discharged from either a PICU or pediatric cardiac ICU, hereafter referred to as ICU, which contributed data to the VPS database from July 1, 2009, to June 30, 2012. Only data from ICUs that managed postoperative pediatric cardiac surgical patients and collected data for both the Pediatric Index of Mortality (PIM)-2 and the Pediatric Risk of Mortality (PRISM)-3 scores were included to assure the availability of multiple physiologic and diagnostic variables to assess for inclusion in a cardiac score. The characteristics of the 55 participating ICUs are shown in **Supplemental Table 1** (Supplemental Digital Content 1, http://links.lww.com/PCC/A179).

Patients

Cardiac surgical patients having surgery either before or after admission to the ICU were selected. Patients unable to have a STAT (10, 11) score derived from the index procedure were excluded. Admission time with respect to cardiac surgery (ATrS) was analyzed due to the hypothesis that physiologic variables are different for patients admitted to a cardiac ICU directly from the operating room as compared with those admitted preoperatively. Neonates who went to the operating room directly from a neonatal ICU were not considered to be admitted to an ICU preoperatively and their postoperative physiologic data were included in the PICSIM model.

Variables

Physiologic variables that were prospectively collected within the first hour and 12 hours of ICU admission and were used to calculate PIM-2 and PRISM-3, respectively, were considered candidate variables. In addition, patient characteristics at admission such as gender, patient origin, age in months, and whether the index procedure was performed before or after admission to the ICU were evaluated. There were 30 physiologic and patient characteristics variables considered originally. In addition, two diagnosis-based criteria (also used in PIM-2) were selected based on their clinical relevance for cardiac surgical patients. This amounted to an initial set of 32 independent variables. Some of the continuous variables were discretized depending on specific thresholds to 1 if YES and

0 if NO, for example, patients having creatinine greater than 0.6 mg/dL = 1; 0 if up to 0.6 mg/dL.

Statistical Methods

Multivariate logistic regression with stepwise selection was used to develop a model to predict mortality in the ICU for cardiac surgical patients. The variables were chosen based on Akaike information criterion and the Mallows *Cp* (12, 13). Thirteen variables, described in **Table 1**, were finally selected for determination of the risk of mortality algorithm. The regression coefficients for the selected variables were determined. Predicted mortality was modeled in the standard fashion:

$$\text{Probability of Death} = \frac{1}{1+\exp(-L)},$$

where the logit (L) is a linear combination of risk factors with the following form:

$$L = b_0 + \sum_{j=1}^{M} b_j \times r_j.$$

The first term on the right side of last equation is an intercept, and the second is a sum of the contributions from each of the risk factors r_j, appropriately weighted by a coefficient b_j, which quantifies how much each risk factor contributed to the outcome among the final 13 risk variables.

The initial cohort was randomly separated into development and validation sets. Seventy-five percent of the initial 16,574 patients were used to determine the coefficients of the risk variables in the logit for mortality, L, by maximization of the likelihood function. The area under the curve (AUC) was compared with STAT score, STAT categories, and RACHS-1 scores. The remaining 25% of the cohort was then used to validate the model by analyzing the receiver operating curve (ROC) and the AUC (14). Calibration of the PICSIM model was tested by the Hosmer-Lemeshow goodness-of-fit test (15). The statistical analysis was done with the R software version 2.15.3 (Vienna, Austria) (16). The AUC was compared with STAT scores, STAT categories, and RACHS-1 scores.

RESULTS

From a total of 123,359 ICU patients discharged from July 2009 to June 2012 in the VPS dataset, there were 16,574 cardiac surgical patients from 55 PICUs. The median number of patients included from each PICU was 188 ranging from 19 to 1,374. Fifty-five percent of the patients were male, 60% were Caucasian, and most of the patients (80%) were admitted to the ICU directly from the operating room. The median age was 7.8 months (range of 0.01 and 673.31), and 2.9% of patients were over 18 years. Of the 16,574 patients, 428 patients (2.6%) died in the ICU.

The results for the validation set (n = 4,143) are summarized in **Table 2**. The SMR values suggest that PICSIM (0.92) predicted deaths for the cardiac surgical population better than PIM-2 (0.54) or PRISM-3 (0.84). The discrimination based on the AUC for PICSIM (0.87) was slightly higher than the others (0.81 for PIM-2 and 0.82 for PRISM-3). **Figure 1** shows a comparison of the ROC curves for the three models, PICSIM (*solid line*), PRISM-3 (*dotted line*), and PIM-2 (*dashed line*) for the validation set. **Table 3** shows the Hosmer-Lemeshow goodness-of-fit calibration of the PICSIM model. **Supplemental**

TABLE 1. Variables Included in the Pediatric Index of Cardiac Surgical Intensive Care Mortality

Variable	Value	Type
Extracorporeal membrane oxygenation within 12 hr of surgery	1 if yes; 0 if no	Categorical
STAT score	0.1–5	Continuous
Hypoplastic left heart syndrome present	1 if yes; 0 if no	Categorical
Mechanical ventilation during the first hour in ICU	1 if yes; 0 if no	Categorical
$FIO_2 > 0.80$	1 if yes; 0 if no	Categorical
Creatinine > 0.60 mg/dL	1 if yes; 0 if no	Categorical
Abnormal hemoglobin < 6 g/dL or > 15 g/dL	1 if yes; 0 if no	Categorical
CO_2 partial pressure greater than 55 mm Hg (from arterial blood gas)	1 if yes; 0 if no	Categorical
Abnormal sodium Na < 137 mmo/L or > 147 mmol/L	1 if yes; 0 if no	Categorical
Patient's average respiratory rate (breaths/min) (range)	7–118	Continuous
Average systolic blood pressure (mm Hg) (range)	31–173	Continuous
STAT score squared	0–25	Continuous
Admission time with respect to cardiac surgery	1 if preoperative; 0 everything else	Categorical

STAT = Society of Thoracic Surgeons-European Association of Cardio-Thoracic Surgery Congenital Heart Surgery Mortality.

TABLE 2. Standardized Mortality Ratio and Area Under the Receiver Operating Characteristic Curve for the Validation Dataset (*n* = 4,143)

Severity of Illness Score	Standardized Mortality Ratio	Area Under the Curve
Pediatric Index of Cardiac Surgical Intensive Care Mortality	0.92	0.87
Pediatric Index of Mortality-2	0.54	0.81
Pediatric Risk of Mortality-3	0.84	0.82
STAT score	–	0.77
STAT category	–	0.75

STAT = Society of Thoracic Surgeons-European Association of Cardio-Thoracic Surgery Congenital Heart Surgery Mortality.

Dashes indicate standardized mortality ratios could not be obtained for STAT score and STAT category as they don't yield mortality predictions in the same manner as the other tools.

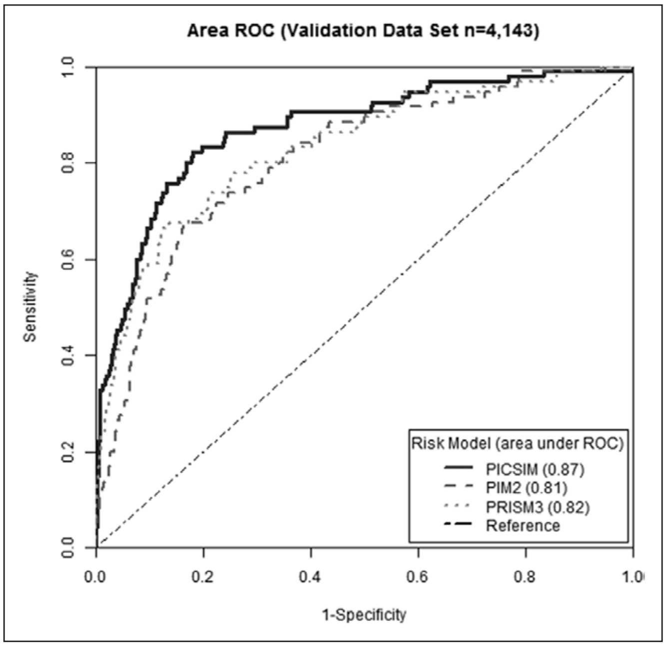

Figure 1. Receiver operating characteristic (ROC) curve for Pediatric Index of Cardiac Surgical Intensive Care Mortality (PICSIM) versus Pediatric Index of Mortality (PIM)-2 and Pediatric Risk of Mortality (PRISM)-3 using the validation dataset (*n* = 4,143). Notice that PICSIM has better discrimination powers than the other models.

Table 2 (Supplemental Digital Content 2, http://links.lww.com/PCC/A180) demonstrates that calibration was acceptable over all age ranges. In addition, we investigated the performance of the PICSIM model in both ATrS (**Table 4**) (ROC values of 0.87 and 0.75 for the postoperative and preoperative admissions, respectively.)

Finally, we compared the performance of the PICSIM score with that of STAT and RACHS-1 (17) scores (**Fig. 2**; **Table 2**; and **Supplemental Table 3**, Supplemental Digital Content 3, http://links.lww.com/PCC/A181). The discrimination based on the AUC for PICSIM score (0.87) was higher compared with the STAT score as a continuous variable (0.77), the STAT categories (0.75), and RACHS-1 (0.74). However, the RACHS-1 analysis was performed with a slightly smaller number of cases (457 less), as over 11% of the cases for which PICSIM scores were able to be determined were unable to be assigned a RACHS-1 score.

The SMRs of the participating institutions ranged from 0 (no deaths) to 3 (three times more deaths observed than predicted). A useful way of comparing SMRs among institutions is the funnel plot (**Fig. 3**). This graphical representation of SMR versus volume per unit is a useful format for understanding the volume outcome relations and for assessing performance outliers while controlling for volume among institutions (18).

DISCUSSION

This study was undertaken to develop an ICU risk of mortality score specifically for cardiac surgical children. Many have suggested that severity-of-illness scores would perform better in homogenous patient populations and have demonstrated that scores developed for the entire critical care population may not perform as well in population subsets, such as cardiac surgical patients (4, 5, 7–9).

Several models have been developed to predict pediatric cardiac surgical mortality based on operative complexity or anatomy, such as RACHS-1 score (17), Aristotle Complexity Score (19), and the STAT score (10, 11). Nonetheless, none of these scores assess the physiologic condition and severity of illness at the time of admission to the ICU. PICSIM is the first attempt to combine physiologic, anatomic, and procedural variables available at the time of ICU admission to predict ICU mortality. The PICSIM variables include indicators of cardiac anatomy and risk (STAT score, hypoplastic left heart syndrome), cardiorespiratory function (respiratory rate, blood pressure, FIO_2, and mechanical ventilation), renal function (creatinine), and laboratory tests (sodium and hemoglobin). Thus, PICSIM includes variables relevant to multiple systems as well as cardiac surgical risk. In addition, the ATrS variable allowed the timing of surgery to be considered. This inclusion of physiologic, cardiac diagnostic, and procedure information found in the aforementioned congenital heart surgery scores may better predict ICU mortality.

As evidenced by improved SMRs and the AUC of the ROC, PICSIM demonstrated improved utility and better discrimination compared with PIM-2 and PRISM-3. This improved performance may be due to the inclusion of variables not included in either PRISM-3 or PIM-2, notably the STAT score and ATrS. A further reason for improved performance may be that it was calibrated specifically in a cardiac surgical population. In addition, PICSIM was found to have improved discrimination compared with the STAT score, STAT categories, and RACHS (**Fig. 2**). Interestingly, our analyses yielded area under the ROC curve results similar to the initial description of the STAT score and STAT category by O'Brien et al (10). Taken together these results indicate that the PICSIM score is more suitable for predicting ICU mortality and therefore more appropriate for assessing ICU performance for cardiac patients than previous methods.

Another advantage of PICSIM is that it has been developed in a large diverse population from 55 ICUs representing 28 states,

TABLE 3. Hosmer-Lemeshow Goodness-of-Fit Test for the Test Set Stratified by Deciles

Pediatric Index of Cardiac Surgical Intensive Care Mortality Decile	*n*	Probability of Death		Observed		Expected	
		Minimum	Maximum	Survived	Died	Survived	Died
1	414	0.0002	0.0018	413	1	413.5	0.5
2	415	0.0018	0.0028	412	3	414.0	1.0
3	414	0.0028	0.0040	414	0	412.6	1.4
4	414	0.0040	0.0055	414	0	412.1	1.9
5	415	0.0055	0.0076	412	3	412.3	2.7
6	414	0.0076	0.0109	412	2	410.2	3.8
7	414	0.0109	0.0157	411	3	408.6	5.4
8	414	0.0157	0.0268	408	6	405.5	8.5
9	415	0.0268	0.0540	398	17	399.5	15.5
10	414	0.0540	0.8404	354	60	351.9	62.1
Total	4,143	0.0002	0.8404	4,048	95	4,040.2	102.8

Chi-square = 11.1; $p = 0.20$. p value of the fit is 0.2, which shows the model calibrates appropriately.

TABLE 4. Discrimination and Goodness-of-Fit Validation for the Preoperative and Postoperative (Admission Time With Respect to Cardiac Surgery) Subsets in the Validation Dataset

Admission Time With Respect to Cardiac Surgery	*n*	Area Under the Curve	Hosmer-Lemeshow, *p*
PICSIM preoperative	746	0.75	0.49
PICSIM postoperative	3,397	0.87	0.22

PICSIM = Pediatric Index of Cardiac Surgical Intensive Care Mortality.

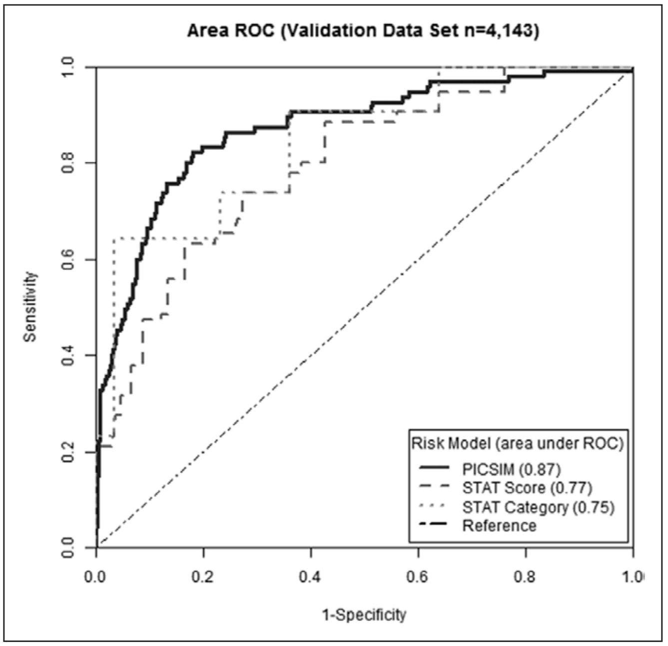

Figure 2. Receiver operating characteristic (ROC) curve for Pediatric Index of Cardiac Surgical Intensive Care Mortality (PICSIM) versus Society of Thoracic Surgeons-European Association of Cardio-Thoracic Surgery Congenital Heart Surgery Mortality (STAT) score and STAT categories using the validation dataset ($n = 4{,}143$).

including patients from small to large ICUs, with or without fellowship programs, general and pediatric free-standing hospitals, demographically and age diverse, thus ensuring wide applicability in the cardiac surgical population.

There are some limitations to this study. First, PICSIM needs to be validated over time for continued reproducibility and could be subject to the same "drift" (20) reported in other prediction models. Second, only variables available in the VPS dataset were used—potentially prospective collection of other variable could improve performance of PICSIM. Third, the model was validated in a cohort of U.S. hospitals and should be validated internationally. Fourth, only those VPS centers that collected PRISM-3 variables were included in the study, which limited the number of eligible sites and may have introduced selection bias. Fifth, although the data used to develop this model were obtained from a high-quality clinical dataset, the potential for misclassification bias remains. Sixth, discrimination and calibration for PICSIM in the preoperative population were rather modest compared with the postoperative population (AUCs, 0.75 and 0.87, respectively) (**Table 4**). These two populations are fundamentally different; the preoperative admission cohort, in addition to not having had surgery, tends to be neonates with greater mortality (5%), whereas the postoperative cohort is older with lower mortality (1.5%). Although inclusion of the ATrS term improved PICSIM performance and permitted one score overall, we advise repeating the PICSIM score following surgery to more accurately reflect the risk of mortality postoperatively. Finally, our approach

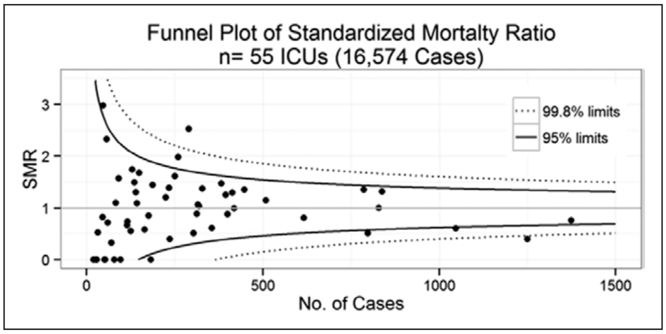

Figure 3. Standard funnel plot showing standardized mortality ratio (SMR) versus volume per unit. It simplifies the determination of outliers when comparing performance among institutions.

used classic regression analysis, similar to the majority of severity-of-illness tools, to develop PICSIM. Newer statistical approaches, such as machine learning (21–24), to discover relationships between multiple variables and outcomes were not explored. Future application of these newer promising "big data" (25) approaches may provide better understanding of severity-of-illness scoring in pediatric critical care.

Severity-of-illness scores, like PICSIM, are necessary for exploring efficiency and efficacy of ICU care (26, 27). The SMR is a cornerstone in benchmarking ICU quality and requires a prediction of mortality score (28). Benchmarking allows comparison among ICUs and internal tracking of improvements in care over time in a given ICU to establish standards for measuring performance and quality, which cannot be improved without appropriate assessment. The PICSIM score can be used to compare ICU performance for cardiac patients adjusting for differences in predicted mortality in a similar fashion to how PIM and PRISM have been used to improve quality for over 30 years (1, 5, 6, 8). As Angus et al (29) editorialized, "it is inevitable, and perhaps desirable, that scoring and risk prediction systems will increasingly become the judges of our clinical activities." Additionally, severity-of-illness adjustment is necessary to compare study cohorts to assure similar mortality in study groups for comparative effectiveness and other research to improve cardiac critical care or to explore the volume outcomes relationship among ICUs as suggest by the funnel plot (**Fig. 3**). Although not yet explored with PICSIM, Pollack and Getson (30) have demonstrated that daily tracking of individual SOI scores can reduce costs and improve efficiency in PICUs. The congenital cardiac care process begins at birth with diagnosis and continues with cardiology care, surgical evaluation, and correction with cardiopulmonary bypass through intensive care and into postsurgical follow-up. All steps of this process must be evaluated to guide improvement of the quality of the care provided for these children. This implies that at each care transition, accurate specific assessment of severity of illness will allow measurement not only of the physiologic status resulting from the preceding stage but also a baseline to compare outcomes of each stage. Thus, PICSIM provides an immediate postoperative assessment that serves as the baseline for the care provided in the ICU. PICSIM is the first specific cardiac SOI tool combining anatomic and physiologic factors assessed at admission to the ICU to enable these multiple approaches to improving the care provided for cardiac surgical patients in ICUs.

ACKNOWLEDGMENTS

We thank the data collection teams from ICUs participating in the Virtual PICU System (VPS) database for their diligence in collecting complete and accurate data. Casey Lauer from VPS formatted all the data utilized in this article. Finally, we thank Dr. Robinder G. Khemani for his insightful comments about the article.

REFERENCES

1. Pollack MM: Outcome analysis. *In*: Textbook of Pediatric Critical Care. Holbrook PR (Ed). Philadelphia, PA, WB Saunders, 1993, pp 1151–1159
2. Pollack MM, Ruttimann UE, Getson PR: Pediatric risk of mortality (PRISM) score. *Crit Care Med* 1988; 16:1110–1116
3. Pollack MM, Patel KM, Ruttimann UE: PRISM III: An updated Pediatric Risk of Mortality score. *Crit Care Med* 1996; 24:743–752
4. Slater A, Shann F, Pearson G; Paediatric Index of Mortality (PIM) Study Group: PIM2: A revised version of the Paediatric Index of Mortality. *Intensive Care Med* 2003; 29:278–285
5. Wetzel RC, Sachedeva R, Rice TB: Are all ICUs the same? *Paediatr Anaesth* 2011; 21:787–793
6. Gemke RJ, van Vught J: Scoring systems in pediatric intensive care: PRISM III versus PIM. *Intensive Care Med* 2002; 28:204–207
7. Czaja AS, Scanlon MC, Kuhn EM, et al: Performance of the Pediatric Index of Mortality 2 for pediatric cardiac surgery patients. *Pediatr Crit Care Med* 2011; 12:184–189
8. Marcin JP, Pollack MM: Review of the acuity scoring systems for the pediatric intensive care unit and their use in quality improvement. *J Intensive Care Med* 2007; 22:131–140
9. Sachdeva RS, Kuhn EM, Gall CM, et al: A need for the development of diagnosis-specific severity of illness tools for risk adjustment [abstract]. *Crit Care Med* 2008; 36(12 Suppl):A83
10. O'Brien SM, Clarke DR, Jacobs JP, et al: An empirically based tool for analyzing mortality associated with congenital heart surgery. *J Thorac Cardiovasc Surg* 2009; 138:1139–1153
11. Jacobs JP, O'Brien SM, Pasquali SK, et al: Variation in outcomes for risk-stratified pediatric cardiac surgical operations: An analysis of the STS Congenital Heart Surgery Database. *Ann Thorac Surg* 2012; 94:564–571
12. Akaike H: Information theory and an extension of the maximum likelihood principle. *In*: Second International Symposium on Information Theory. Petrov BN, Csaki F (Eds). Budapest, Akademia Kiado, 1973, pp 267–281
13. Mallows CL: Some comments on Cp. *Technometrics* 1973; 15:661
14. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29–36
15. Hosmer DW, Lemeshow D: Assessing the fit of the model. *In*: Applied Logistic Regression. Second Edition. Hosmer DW, Lemeshow S, Cook ED (Eds.). Hoboken, NJ, John Wiley & Sons, 2000; pp 143–202
16. R Development Core Team: R: A Language and Environment for Statistical Computing. Vienna, Austria, Foundation for Statistical Computing, 2013. Available at: http://www.R-project.org. Accessed July 20, 2014
17. Jenkins KJ, Gauvreau K, Newburger JW, et al: Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg* 2002; 123:110–118
18. Spiegelhalter DJ: Funnel plots for comparing institutional performance. *Stat Med* 2005; 24:1185–1202
19. Lacour-Gayet F, Clarke D, Jacobs J, et al; Aristotle Committee: The Aristotle score: A complexity-adjusted method to evaluate surgical results. *Eur J Cardiothorac Surg* 2004; 25:911–924
20. Wetzel RC, Kuhn EM, Gall CM, et al: Demonstrating time-dependent drift in a validated risk adjustment tool. *Crit Care Med* 2008; 36(12S):A8

21. Song X, Mitnitski A, Cox J, et al: Comparison of machine learning techniques with classical statistical models in predicting health outcomes. *Medinfo* 2004; 11:736–740
22. Jensen PB, Jensen LJ, Brunak S: Mining electronic health records: Towards better research applications and clinical care. *Nat Rev Genet* 2012; 13:395–405
23. Pearce CB, Gunn SR, Ahmed A, et al: Machine learning can improve prediction of severity in acute pancreatitis using admission values of APACHE II score and C-reactive protein. *Pancreatology* 2006; 6:123–131
24. Wu J, Roy J, Stewart WF: Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches. *Medical Care* 2010; 48:S106–S113
25. Raghupathi W, Raghupathi V: Big data analytics in healthcare: Promise and potential. *Health Inf Sci Syst* 2014; 2:3
26. Marcin JP, Pollack MM: Review of the methodologies and applications of scoring systems in neonatal and pediatric intensive care. *Pediat Crit Care Med* 2000; 1:20–27
27. Gaies M, Jeffries HE, Jacobs JP, et al: Measuring quality and outcomes in pediatric cardiac critical care. *Prog Pediatr Cardiol* 2012; 33:33–36
28. Gemke RJ, Bonsel GJ: Comparative assessment of pediatric intensive care: A national multicenter study. Pediatric Intensive Care Assessment of Outcome (PICASSO) Study Group. *Crit Care Med* 1995; 23:238–245
29. Angus DC, Pinsky MR: Risk prediction: Judging the judges. *Intensive Care Med* 1997; 23:363–365
30. Pollack MM, Getson PR: Pediatric critical care cost containment: Combined actuarial and clinical program. *Crit Care Med* 1991; 19:12–20