Morbidity and mortality prediction in pediatric heart surgery: Physiological profiles and surgical complexity

John T. Berger, MD,[a] Richard Holubkov, PhD,[b] Ron Reeder, PhD,[b] David L. Wessel, MD,[a] Kathleen Meert, MD,[c] Robert A. Berg, MD,[d] Michael J. Bell, MD,[e] Robert Tamburro, MD, MSc,[f] J. Michael Dean, MD,[c] and Murray M. Pollack, MD,[a,g] for the Eunice Kennedy Shriver National Institute of Child Health and Human Development Collaborative Pediatric Critical Care Research Network

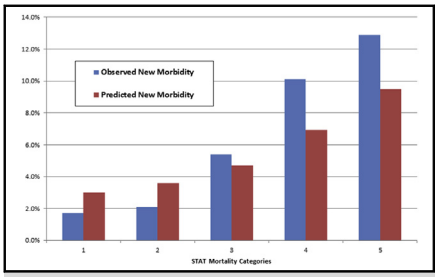
ABSTRACT

Objectives: Outcome prediction for pediatric heart surgery has focused on mortality but mortality has been significantly reduced over the past 2 decades. Clinical care practices now emphasize reducing morbidity. Physiology-based profiles assessed by the Pediatric Risk of Mortality (PRISM) score are associated with new significant functional morbidity detected at hospital discharge. Our aims were to assess the relationship between new functional morbidity and surgical risk categories (Risk Adjustment for Congenital Heart Surgery [RACHS] and Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk [STAT]), measure the performance of 3-level (intact survival, survival with new functional morbidity, or death) and 2-level (survival or death) PRISM prediction algorithms, and assess whether including RACHS or STAT complexity categories improves the PRISM predictive performance.

Methods: Patients (newborn to age 18 years) were randomly selected from 7 sites (December 2011-April 2013). Morbidity (using the Functional Status Scale) and mortality were assessed at hospital discharge. The most recently published PRISM algorithms were tested for goodness of fit, and discrimination with and without the RACHS and STAT complexity categories.

Results: The mortality rate in the 1550 patients was 3.2%. Significant new functional morbidity rate occurred in 4.8%, increasing from 1.8% to 13.9%, 1.7%, and 12.9% from the lowest to the highest RACHS and STAT categories, respectively. The 3-level and 2-level PRISM models had satisfactory goodness of fit and substantial discriminative ability. Inclusion of RACHS and STAT complexity categories did not improve model performance.

Conclusions: Both mortality and new, functional morbidity are important outcomes associated with surgical complexity and can be predicted using PRISM algorithms. Adding surgical complexity to the physiologic profiles does not improve predictor performance. (J Thorac Cardiovasc Surg 2017;154:620-8)



New functional status morbidity increases with Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk score mortality categories and can be accurately predicted.

Central Message

New, functional morbidity is associated with surgical complexity and can be predicted with mortality by a physiology-based algorithm.

Perspective

Mortality is infrequent, whereas new functional morbidity at hospital discharge is common, after congenital heart surgery. Studies focused on mortality may miss meaningful clinical issues and require large samples. We found that new functional morbidity at hospital discharge as well as mortality increased with increasing surgical risk and can be simultaneously predicted by a physiology-based algorithm.

See Editorial Commentary page 629.

See Editorial page 618.

Outcome prediction for critically ill children following congenital heart surgery has centered on operative mortality. One prominent approach uses the anatomic diagnosis and/or specific operation performed for palliation or repair as the core risk-adjustment methodology. The Risk Adjustment for Congenital Heart Surgery (RACHS) score relies

From the Departments of [a]Pediatrics, Children's National Medical Center; [b]Pediatrics, University of Utah School of Medicine, Salt Lake City, Utah; [c]Pediatrics, Children's Hospital of Michigan, Detroit, Mich; [d]Pediatrics and [e]Critical Care Medicine, Children's Hospital of Philadelphia, Philadelphia, Pa; [f]Pediatric Trauma and Critical Illness Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institutes of Health, Bethesda, Md; and [g]George Washington University School of Medicine and Health Sciences, Washington, DC.

Supported, in part, by cooperative agreements from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services (grants U10HD050096, U10HD049981, U10HD049983, U10HD050012, U10HD063108, U10HD063106, U10HD063114 and U01HD049934). This content is solely the responsibility of the authors and does not necessarily represent the views of the National Institutes of Health.

M.M.P. and R.H. had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Received for publication May 2, 2016; revisions received Dec 15, 2016; accepted for publication Jan 8, 2017; available ahead of print March 6, 2017.

Address for reprints: Murray M. Pollack, MD, Children's National Medical Center, 111 Michigan Ave, NW, Washington, DC 20010 (E-mail: mpollack@childrensnational.org).

0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2017.01.050

Abbreviations and Acronyms

AUC = area under the curve
FSS = Functional Status Scale
ICU = Intensive care unit
PICSIM = Pediatric Index of Cardiac Surgical Intensive Care Mortality
PRISM = Pediatric Risk of Mortality
RACHS = Risk Adjustment for Congenital Heart Surgery
ROC = receiver operating characteristic
STAT = Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk
STS-CHSD = Society for Thoracic Surgery Congenital Heart Surgery Database
TOPICC = Trichotomous Outcome Prediction in Critical Care
VUS = volume under the surface

Scanning this QR code will take you to the supplemental tables and video for this article.

on subjective assessments of operative risk and cardiac anatomy by congenital heart surgeons and pediatric cardiologists.[1] The most recent method, the 2014 Society for Thoracic Surgery Congenital Heart Surgery Database (STS-CHSD) Mortality Risk (STAT) model, estimates risk by calculating an expected rate of mortality that accounts for the operation performed and a number of preoperative variables.[2,3] Mortality risks for individuals are computed using the risk of each combination of primary procedure, age group, and other cofactors to adjust for individual patient factors. Recently, these cofactors have expanded to include preoperative intensive care unit (ICU) clinical factors and therapies.[4] The risk for inpatient morbidity has been similarly developed.[5] This approach is the foundation for a major quality program.[3,6]

Physiology-based severity of illness methods used in adult, pediatric, and neonatal intensive care for decades have also centered on mortality.[7-10] The Pediatric Risk of Mortality (PRISM) score is a frequently used, physiology-based measure that assigns numeric values reflective of mortality risk to derangements of 17 commonly measured physiologic variables. The PRISM score is the summation of these values, whereas mortality risk is computed using the PRISM score and other cofactors.[8] The numeric PRISM score is termed severity of illness.[11] PRISM has been a foundation of national quality programs. It has performed well in congenital heart surgery patients consistent with the observation that postprocedure physiological status reflects mortality risk.[8] Recently, PRISM has undergone a revision of its data collection methods.[12,13] Most importantly, the PRISM outcome algorithm estimates simultaneously the risk of new functional morbidity as well as mortality at hospital discharge.[13] PRISM algorithms are also available for estimation of mortality risk alone.[12] PRISM prediction algorithms have not been rigorously assessed in a modern cohort of congenital heart surgery patients.

A third approach for pediatric risk assessment is based on general and targeted categorical variables, and a limited set of physiologic variables and therapies. The Pediatric Index of Cardiac Surgical Intensive Care Mortality (PICSIM)[14] overlaps with the Pediatric Index of Mortality, which did not perform well in cardiac surgery patients.[15,16] Because most of the PICSIM predictive power comes from the surgical complexity score, its use to assess intensive care quality is limited.[17]

Mortality rates in pediatric heart surgery and critical care are low and decreasing, with rates reported to be <4%.[2,14,18] Yet, modern risk assessment methods continue to focus on operative or intensive care mortality. In contrast, new morbidity rates assessed as functional status changes in critically ill children measured at hospital discharge are approximately twice as high as mortality rates and it has been suggested that functional morbidity is replacing mortality.[19] Recently, the Eunice Kennedy Shriver National Institute of Child Health and Human Development Collaborative Pediatric Critical Care Research Network developed a granular measure of functional morbidity that is age independent and sufficiently rapid, accurate, and reliable for population-based outcome studies.[20] This method, the Functional Status Scale (FSS), is a significant improvement over common subjective scales.[21,22] Importantly, we recently demonstrated that the development of new functional status morbidities was associated with physiological status early in the ICU course in a manner that parallels the association between physiological status and mortality. Further, we demonstrated that we could simultaneously estimate the risk of both functional morbidity and mortality from data obtained during the first 4 hours of intensive care.[13]

The analyses described in this article had 3 specific aims. Our first aim was to examine how the risk of developing new, significant functional morbidity was associated with levels of a physiology-based score, and with the risk categories of the RACHS and STAT scores. Second, we assessed the performance of the recently published 3-level PRISM prediction algorithms (ie, death; survival with new, significant functional morbidity; and survival without new, significant

functional morbidity [intact survival]) and 2-level prediction algorithm (ie, survival or death) in a contemporary sample of pediatric heart surgery patients.[13] This assessment included the performance of an objective algorithm to determine the PRISM observation time because some patients are admitted preoperatively. Third, we assessed the potential for prediction improvement by including the risk categories from RACHS and STAT and other cardiac descriptors in the PRISM prediction equations.

METHODS

This investigation used the cardiovascular surgery patients in the Trichotomous Outcome Prediction in Critical Care (TOPICC) database collected by the Child Health and Human Development Collaborative Pediatric Critical Care Research Network. Detailed methods of TOPICC data collection have been previously described.[13] The central aim of TOPICC was to assess the relationship between physiologic profiles and the development of functional morbidity. In brief, there were 7 sites with 1 site composed of 2 institutions. Randomly selected patients, newborn to younger than age 18 years, admitted to participating pediatric and cardiac ICUs from December 4, 2011, to April 7, 2013, were included for analysis and stratified by hospital.[13] Moribund patients (ie, vital signs incompatible with life for the first 2 hours after ICU admission) were excluded. Only the first ICU admission during a hospitalization was included. Demographic data were obtained on admission. All participating institutional review boards approved the protocol. Detailed institutional data along with other analyses have been published.[13,19,21,23,24] For additional details concerning patient and site-level data, outcomes, and physiologic data see Appendix E1.

Outcomes

Functional morbidity, mortality, and survival without new functional morbidity were assessed at hospital discharge. New morbidity affecting a significant decrement in functional status was assessed with the FSS for the baseline status (before the acute illness requiring ICU admission) and at hospital discharge. The FSS is an age-independent assessment of functional status that can be determined from the medical record or from health care providers' input.[20] It was developed as a granular and objective instrument suitable for large pediatric outcome studies. The 6 domains (mental status, sensory, communication, motor function, feeding, and respiratory) are individually scored with a range from 1 (normal) to 5 (very severe dysfunction). The operational definitions and manual for the classifications have been published.[20] Newborns never achieving a stable baseline are assigned a FSS score of 6; this was operationalized by assigning a FSS of 6 to patients admitted to the study sites who were aged 0 to 2 days and to transfers from another facility who were aged 3 to 6 days. New morbidity was defined as an increase in the FSS score $\geq$3 points from baseline to hospital discharge; changes of this magnitude indicate substantial worsening of functioning. Previous analysis indicated that more than 95% of these children had a change of 2 or more points in a single domain, a clearly significant functional change. Functional morbidity occurs in essentially all ages and types of patients, in relatively equal proportions, and involves all FSS domains.[19]

Measurement of Physiological Status

Physiological status was measured with the PRISM score with a shortened time interval (2 hours before admission to 4 hours after admission for laboratory data and the first 4 hours of ICU care for other physiological variables). Outcome prediction using this time interval included separation of the total PRISM into neurologic and nonneurologic components and other patient factors.[12,13]

Congenital Cardiac Conditions

Only cardiovascular surgery patients were included in this analysis. Classifications as 1 or 2 ventricle, cyanotic or acyanotic, and by the RACHS and STAT categories, were done by a cardiologist (J.T.B.) based on the anatomic diagnosis and operative procedure in the operative report and the admission diagnostic information and blinded to the outcomes.[18,25] Operations involving combinations of procedures were assigned to the procedure with the highest mortality category. Cyanosis was based on preoperative anatomy and description from the surgical notes. Patients were classified as single- or 2-ventricle repair based on evidence of ventricular hypoplasia using the type of operation and operative report.

The time interval for assessing PRISM data was modified for cardiac patients younger than age 91 days because some institutions admit young infants to the ICU before a cardiac intervention to optimize clinical status, and not for intensive care; in these cases, the postintervention period more accurately reflects intensive care. However, in other infants for whom the cardiac intervention is delayed after ICU admission, the intervention is a therapy required due to failed medical management of the acute condition; in these infants, the routine PRISM data collection time interval is an appropriate reflection of critical illness. A priori, we identified infants for whom it would be more appropriate to utilize data from the 4 hours after the cardiac intervention (postintervention time interval) and those for whom using the admission time interval was more appropriate and operationalized this decision on the conditions likely to present within the first 90 days, the time period when the vast majority of these conditions present. This approach has been detailed elsewhere (Table E1).[12,13] We assessed the adequacy of fit, as well as performance, of the PRISM prediction models in the age groups of 90 days or younger and older than 90 days using standardized morbidity and mortality ratios.

Statistical Methods

Statistical analyses utilized SAS 9.4 (SAS Institute Inc, Cary, NC) for descriptive statistics, model development, and fit assessment, and R 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria) for evaluation of predictive ability. The statistical analysis was under the direction of R.H.

Patient characteristics were descriptively compared and evaluated across sites using the Kruskal-Wallis test for continuous variables, and Fisher exact test for categorical variables.

Predicted numbers of events were calculated using probabilities from previously published models constructed from the TOPICC cohort and these calculated probabilities were used to determine the predicted outcomes in the analyses.[12,13] Goodness of fit of these models was assessed using the Hosmer-Lemeshow test for logistic models and an extension to 3 outcomes.[26] We treated the cardiac cohort as an independent sample in terms of applicable test degrees of freedom, because this cohort is a small subset of the entire population defined per clinical criteria, and includes validation set cases not used in the TOPICC model construction. To maintain the validity of the Hosmer-Lemeshow test (for which expected event counts should be $\geq$5 within most evaluated cells), subjects were sorted in order of increasing predicted probability of mortality, and then divided into risk categories each containing approximately 7 expected deaths. Reported goodness-of-fit findings were robust to alternate risk category specifications.[26] Reported goodness-of-fit findings were robust to the number of such categories used. For reported standardized mortality ratios, the Breslow-Day method was used to calculate 2-sided 95% confidence intervals.

Discrimination was assessed by 2-dimensional receiver operating characteristic curves for the survival/death model and by 3-dimensional volume under the surface (VUS) for the 3-level outcome. Two-dimensional ROC curves were generated, with area under the curve (AUC) calculated and its variability estimated, using the SAS logistic procedure. VUS for discriminating between the 3 outcomes is reported using the R_H triplet-classification rule of Mossman.[27] The VUS has a value of one-sixth under a model with no discriminatory ability; we also report the average

dichotomized C-index (the average of the areas under the curve considered over all possible ordered dichotomizations of the outcome, the value of which with no model discrimination is 0.5) as an alternate summary measure of multidimensional model discriminatory ability.

For assessing whether adding a cardiac measure (RACHS, STAT, single- vs 2-ventricle anatomy, or cyanotic vs acyanotic status) improved the predictive ability of the published PRISM models, the cardiac measure was added as a categorical predictor to a logistic model (dichotomous or trichotomous) that held each patient's PRISM predicted outcome probabilities fixed using an offset term. The STAT mortality categories were added to our model without the use of additional preoperative patient characteristics. This modeling used SAS PROC NLMIXED. Significance of improvement for a model including a cardiac-measure predictor was assessed by comparing its likelihood value to that of the published PRISM model applied to this population. We also quantified potential improvement in discrimination via the AUC and VUS.

RESULTS

The overall sample contained 10,078 patients, of whom 1550 underwent a cardiac surgery. Sample characteristics at the site level and overall are shown in Table E2, including age, age distribution, STAT categories, ICU, and hospital lengths of stay, PRISM scores, outcomes, and the classifications of cyanotic or acyanotic and single- or 2-ventricle anatomy. Of the cardiac interventions, 1199 (77.4%) had 2-ventricle anatomy and 351 (22.6%) were single-ventricle patients. A total of 871 (56.2%) were acyanotic and 679 (43.8%) were cyanotic. Based on information available for the interventions performed, the RACHS score was calculable in 1447 of these cardiac patients, whereas the STATS categorization was achievable in 1534 patients. Overall, the mortality rate was 3.2% and the new functional morbidity rate was 4.8%.

The new functional morbidity and mortality rates for each RACHS and STAT category are displayed in Table 1 and illustrated for the STAT categories in Figure 1. Overall, both the observed and predicted functional morbidity and mortality rates significantly increased with increasing RACHS and STAT categories. The only exception was the RACHS 5 category, which had too few cases for statistical stability. In particular, the new functional morbidity rates increased from 1.8% to 13.9% and 1.7% and 12.9% from the lowest to the highest severity categories for RACHS and STAT, respectively.

Next, we tested the performance of the PRISM 3-level prediction model predicting intact survival, new functional morbidity at hospital discharge, and death. Initially, we assessed the performance of the PRISM prediction models in those younger than age 90 days and those older than age 90 days. The standardized morbidity and mortality ratios performed well, indicating the decision matrix for assigning the PRISM observation period was sufficient (Table E3). In assessing the model performance, we first used the categories of RACHS (combining levels 1 with 2 and 5 with 6 due to small numbers of within-cell events) and STAT for the severity categories for the goodness-of-fit risk groups. Both RACHS and STAT (Table 1) demonstrated acceptable fit (RACHS: $\chi^2 = 6.972$; df $= 8$; $P = .540$; STAT: $\chi^2 = 13.558$; df $= 10$; $P = .19$) Next, we used 7 risk categories constructed with at least 7 expected mortalities in each cell to assess the goodness of fit for the intact survival/new morbidity/death (Table 2) and survival/death models (Table 3). Overall, for the 3-level model, 49.8 deaths

TABLE 1. Observed and predicted mortality and new functional morbidity in RACHS and STAT categories. Both observed and predicted mortality and functional morbidity rates increased with increasing severity categories for both systems (both *P* values < .0001). Predicted new functional morbidity rates also increased with increasing severity categories (RACHS, *P* = .0032; STAT, *P* = .0009)

	N	Median age (mo)	Crude mortality	Predicted mortality	Mortality	Crude new morbidity	Predicted morbidity	Morbidity
RACHS								
1	114	55	0 (0.0)	0.7 (0.6)	0 (NA-5.3)	2 (1.8)	2.9 (2.5)	0.7 (0.1-2.5)
2	585	6	10 (1.7)	9.2 (1.6)	1.1 (0.5-2.0)	15 (2.6)	21.7 (3.7)	0.7 (0.4-1.1)
3	517	9	17 (3.3)	16.2 (3.1)	1.1 (0.6-1.7)	24 (4.6)	24.1 (4.7)	1.0 (0.6-1.5)
4	149	0	8 (5.4)	11.5 (7.7)	0.7 (0.3-1.4)	15 (10.1)	10.5 (7.1)	1.4 (0.8-2.4)
5	3	0	0 (0.0)	0.3 (11.2)	0 (NA-10.9)	0 (0.0)	0.3 (9.7)	0 (NA-12.2)
6	79	0	9 (11.4)	10.4 (13.1)	0.9 (0.4-1.7)	11 (13.9)	7.7 (9.7)	1.4 (0.7-2.6)
Unable to classify	103	69	6 (5.8)	1.5 (1.5)	3.9 (1.4-8.5)	7 (6.8)	3.8 (3.7)	1.9 (0.7-3.8)
STAT								
1	423	28	2 (0.5)	3.8 (0.9)	0.5 (0.1-1.9)	7 (1.7)	12.7 (3.0)	0.5 (0.2-1.1)
2	513	15	10 (1.9)	8.2 (1.6)	1.2 (0.6-2.2)	11 (2.1)	18.5 (3.6)	0.6 (0.3-1.1)
3	205	5	6 (2.9)	6 (2.9)	1.0 (0.4-2.2)	11 (5.4)	9.7 (4.7)	1.1 (0.6-2.0)
4	308	0	21 (6.8)	20.6 (6.7)	1.0 (0.6-1.6)	31 (10.1)	21.1 (6.9)	1.5 (0.99-2.1)
5	85	0	9 (10.6)	10.7 (12.6)	0.8 (0.4-1.6)	11 (12.9)	8.1 (9.5)	1.4 (0.7-2.4)
Unable to classify	16	40	2 (12.5)	0.5 (2.8)	4.4 (0.5-15.8)	3 (18.8)	0.8 (5.2)	3.6 (0.7-10.6)

Values are presented as n (%), or standardized morbidity/mortality ratio (95% confidence interval) unless otherwise noted. *RACHS*, Risk Adjustment for Congenital Heart Surgery; *STAT*, Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk; *NA*, not available.

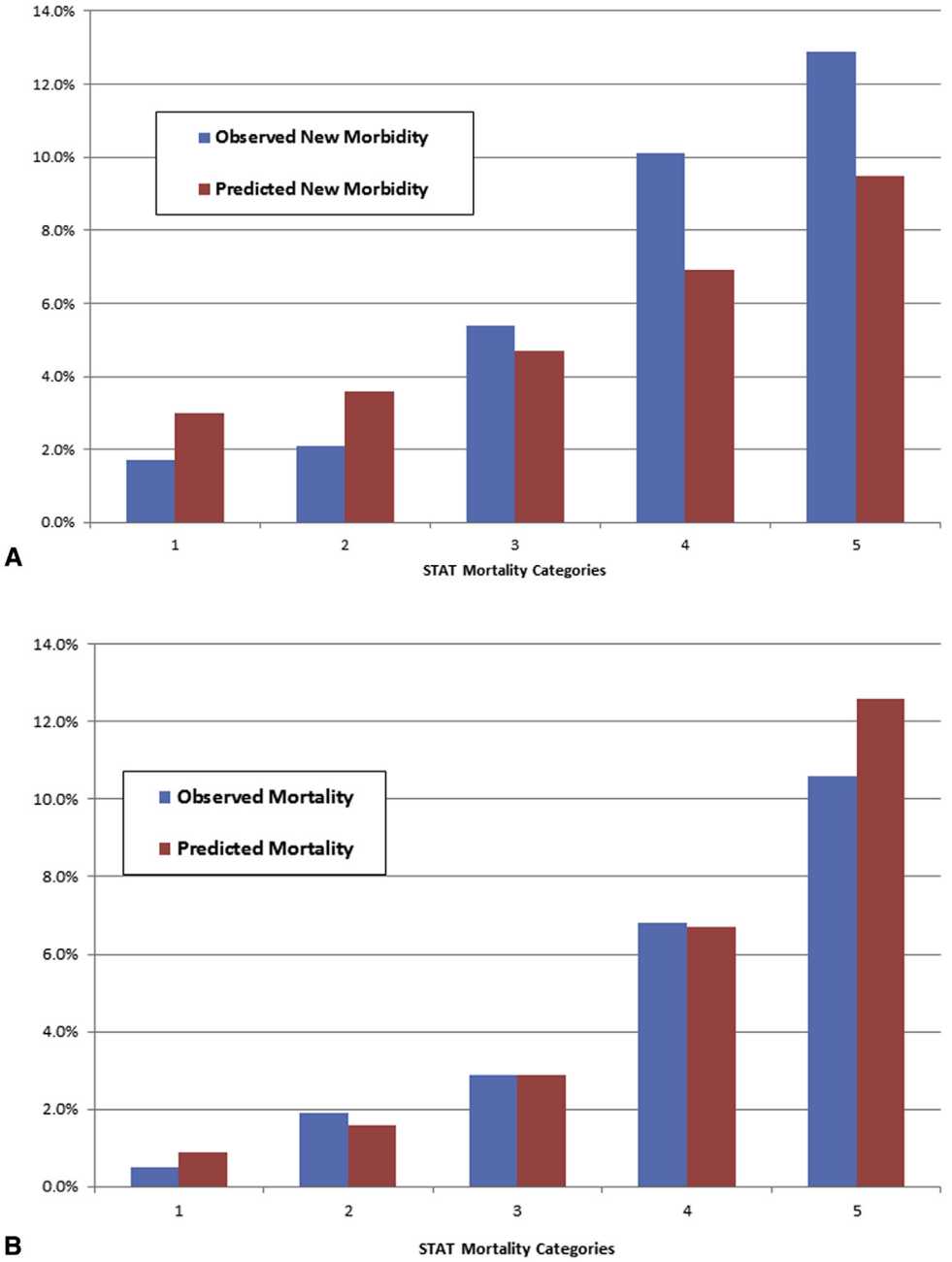


FIGURE 1. A, Observed and predicted new functional morbidity and (B) Observed and predicted mortality for Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk (*STAT*) score mortality categories. Both observed and predicted functional morbidity and mortality rates increased with increasing STAT mortality categories ($P < .0001$) (see Table 1 for details.)

were predicted and 50 were observed (standardized mortality ratio = 1.0) and 71.0 new functional morbidities were predicted and 74 were observed (standardized morbidity ratio, 0.96). The goodness of fit was acceptable ($P = .31$). Discriminative ability was excellent, with a VUS of 0.46 (vs a chance value of 0.17). The average dichotomized C-index for this population was 0.82. For the dichotomous model, 50.1 deaths were expected and 50 were observed (standardized mortality ratio = 0.86). The goodness of fit was acceptable ($P = .474$). The AUC of the survival/death model was 0.83.

The standardized mortality and morbidity ratios of the dichotomous and trichotomous predictors in the clinical categories of cyanotic or acyanotic and single- or 2-ventricle lesions are shown in Table 4. The prediction performance based on standardized mortality ratios was

TABLE 2. Goodness of fit test for the new functional morbidity-intact survival-death model. The Hosmer-Lemeshow χ^2 test statistic = 16.036 (df = 14; P = .31). The volume under the surface was 0.46 (chance = 0.17). The SMRs were 1.00 and 0.96, respectively

	Deaths			Morbidity		
Risk group	E	O	SMR (95% confidence interval)	E	O	SMR (95% confidence interval)
0	7	13	1.9 (1.0-3.2)	30	24	0.8 (0.5-1.2)
1	7	5	0.7 (0.2-1.7)	14.3	20	1.4 (0.9-2.2)
2	7	5	0.7 (0.2-1.7)	9.2	12	1.3 (0.7-2.3)
3	7	8	1.1 (0.5-2.2)	6.8	7	1.0 (0.4-2.1)
4	7	6	0.9 (0.3-1.9)	5.2	7	1.3 (0.5-2.8)
5	7.4	8	1.1 (0.5-2.1)	3.8	3	0.8 (0.2-2.3)
6	7.3	5	0.7 (0.2-1.6)	1.7	1	0.6 (0.0-3.2)
Total	49.8	50	1.0 (0.7-1.3)	71	74	1.0 (0.8, 1.3)

E, Expected; *O*, observed; *SMR*, standardized morbidity/mortality ratio.

acceptable in all groups. Finally, we assessed the potential improvement in model performance by separately adding the RACHS categories, STAT categories, cyanotic/acyanotic factor, and single- or 2-ventricle factor to the PRISM prediction models. Table 5 displays the significance level for adding each factor, and the improvement in the VUS or AUC if the factor is added. In all cases, inclusion of the factor did not significantly improve the model performance.

DISCUSSION

Mortality from both pediatric heart surgery and pediatric ICUs has fallen to low rates, making mortality an insensitive outcome for care assessments and therapeutic studies without very large samples. Because much of pre- and postoperative care focuses on reducing functional morbidity as well as mortality, functional status is an important outcome. In this pediatric cardiac surgery population, the overall rate of significant, new functional morbidity was 50% higher than mortality; in the general ICU population, this rate is approximately twice as high as mortality. Importantly, the new functional morbidity risk increased more than 3-fold from the lowest to the highest surgical risk categories.

TABLE 3. Goodness-of-fit test for the survival-death model. The Hosmer-Lemeshow χ^2 test = 6.58 (df = 7; P = .474). The area under the curve was 0.83 ± 0.03. The standardized mortality ratio was 1.0

Risk group	E	O	SMR (95% confidence interval)
0	7.0	8	1.1 (0.5-2.2)
1	7.0	11	1.6 (0.8-2.8)
2	7.1	4	0.6 (0.2-1.5)
3	7.1	7	1.0 (0.4-2.0)
4	7.0	7	1.0 (0.4-2.1)
5	7.1	8	1.1 (0.5-2.2)
6	7.8	5	0.6 (0.2-1.5)
Total	50.1	50	1.0 (0.7-1.3)

E, Expected; *O*, observed; *SMR*, standardized morbidity/mortality ratio.

The PRISM models estimating functional morbidity and mortality risk performed well. Discrimination for mortality in these models is similar to the older PRISM models,[14] although the observation time is substantially shorter; hospital outcome is used, which has been harder to predict; only the first ICU admission is included; and the data sampling period is objectively assigned based on age and time to intervention. Importantly, the PRISM methodology was specifically developed to minimize the potential for institutional bias or gaming at the expense of model performance. For example, the observation time was chosen to minimize the potential for institutional care practices to affect the PRISM score,[28] modeling of hospital outcome was specifically chosen instead of ICU outcome to minimize the effect of premature ICU discharge with readmission, and the objective process to determine the sampling time period for heart surgery in infants younger than age 90 days was created to this accommodate intercenter variability.

The discrimination is slightly less than the reported discrimination in the new STS-CHSD model and the PICSIM score.[2,14] The PICSIM score uses postoperative therapies as well as a 12-hour postoperative sampling period for some of the physiologic variables. The use of postoperative therapies in risk models can create bias. Although their inclusion would improve predictor performance, therapies are intentionally not included in the PRISM models because separating physiology from therapy allows independent assessment of the timely and appropriate use of therapy (quality of care).

Importantly, adding surgical complexity classifications to the physiology-based model did not improve model performance, indicating that the physiology-based PRISM score captured most of the information concerning surgical complexity. Because the relationship between functional morbidity and physiological status is sufficiently precise for accurate functional morbidity prediction, we believe that functional morbidity risk as well as mortality risk is reflected in large part through postoperative physiological status. However, we do not have direct confirmation of this

TABLE 4. SMRs for the cyanotic/acyanotic and 1- or 2-ventricle classifications

		Survival-death model			Morbidity-intact survival-death model					
		Mortality			Morbidity			Mortality		
Variable	n	O	E	SMR (95% confidence interval)	O	E	SMR (95% confidence interval)	O	E	SMR (95% confidence interval)
Cyanotic	679	39	35.7	1.1 (0.8-1.5)	45	40.0	1.1 (0.8-1.5)	39	36.0	1.1 (0.8-1.5)
Acyanotic	871	11	14.4	0.8 (0.4-1.4)	29	31.0	0.9 (0.6-1.3)	11	13.7	0.8 (0.4-1.4)
1 Ventricle	351	21	16.7	1.3 (0.8-1.9)	28	20.0	1.4 (0.9-2.0)	21	16.7	1.3 (0.8-1.9)
2 Ventricles	1199	29	33.4	0.9 (0.6-1.2)	46	51.0	0.9 (0.7-1.2)	29	33.0	0.9 (0.6-1.3)

O, Observed; *E*, expected; *SMR*, standardized morbidity/mortality ratio.

causal relationship. Other conditions associated with congenital heart disease could be contributing to discharge functional status.[29-31]

There are 2 general uses for prediction models such as the ones presented in this analysis. First, they can focus on evaluating of systems by adjusting for patient characteristics. Our analyses focused on this use. The advantage to the PRISM models based on postoperative physiologic profiles is that they more directly assess ICU performance. Because the STS-CHSD and PICSIM model performances are based predominantly on the surgical procedure performed,[2,14] they assess risk at the time the patient enters the operating room, whereas PRISM assesses risk when the patient enters the ICU. Methods such as the STS-CHSD better assess the whole system, including the diagnostic assessment, determining the operative approach, surgical and anesthesia operative performance, and pre- and postoperative care. Therefore, the 2 approaches are complementary. We believe that if ICU assessment is paramount, a physiology-based approach is preferable. Second, prediction models potentially can be used at the individual patient level. Our analyses presenting performance and outcomes within subpopulations defined by various risk criteria did not focus on this use.

There are potentially significant limitations to this analysis. First, the sample size is relatively small in comparison to other similar studies. Although the sample size is sufficient to uncover major influences on the PRISM models, it is possible that a larger sample would have uncovered other issues with significant, but weaker influences on the model. Second, it was assumed that newborns had a normal baseline functional status because they never achieved a baseline state other than their in utero condition. Although the PRISM models perform well in all age groups, including neonates and young infants, we have been unable to rigorously test this assumption.

Several challenges remain in this new era of outcome assessment. First, do assessment methods change the quality of care in individual institutions? We lack sufficient evidence that the time and effort spent collecting these data are appropriately used by the participating institutions to improve care. The efforts to ensure reliable methods with relevant outcomes that are unbiased are foundational to evaluating and improving care. Second, we need to better understand the relationship between hospital discharge and long-term outcomes for all types of critically ill patients, including pediatric cardiac surgery patients. Long-term outcomes are an important aspect of the effectiveness of care, but the long observation times make this difficult and challenging. A better understanding of the relationship between short-term and long-term outcomes would enable us to assess and improve short-term outcomes with the security that it would translate into improved long-term outcomes.

CONCLUSIONS

There is strong relationship between new, significant functional morbidity at hospital discharge and surgical

TABLE 5. Significance of adding RACHS, STAT score, cyanotic/acyanotic, and single- or 2-ventricle covariates to the Pediatric Risk of Mortality prediction models

	Morbidity-intact survival-death (trichotomous) model		Survival-death (binary) model	
Factor	Significance level*	VUS with/without factor	Significance level*	AUC with/without factor
RACHS†	.53	0.483/0.497	.78	0.854/0.854
STAT‡	.16	0.472/0.490	.83	0.836/0.842
Cyanotic-acyanotic	.75	0.457/0.467	.50	0.830/0.832
Single-two ventricle	.19	0.457/0.466	.37	0.830/0.832

VUS, Volume under the surface; *AUC*, area under the curve; *RACHS*, Risk Adjustment for Congenital Heart Surgery; *STAT*, Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk. *For the likelihood ratio test, adding the factor to a model with outcome probabilities fit using the published model coefficients. †RACHS categories 1 and 2, and categories 5 and 6, were combined to achieve sufficient numbers of outcomes in category levels to allow model convergence. ‡All 5 STAT categories were used in modeling.

VIDEO 1. Contextual analysis and summary of the results. Video available at: http://www.jtcvsonline.org/article/S0022-5223(17)30200-3/addons.

complexity as well as postoperative physiological status. Because new functional morbidity is an important patient outcome that is substantially more common than mortality, it should be included as an outcome in quality and other studies for children following congenital heart surgery (see Video 1 for a more expanded discussion of this issue).

Conflict of Interest Statement

There is a patent pending regarding the trichotomous outcome predictor by Children's National Health System, the employer of Drs Pollack, Berger, and Wessel. All other authors have nothing to disclose with regard to commercial support.

The authors thank the following individuals for their contributions: Teresa Liu, MPH, CCRP, University of Utah (project management, Data Coordinating Center); Jean Reardon, MA, BSN, RN, Children's National Medical Center (institutional project management, data collection); Elyse Tomanio, BSN, RN, Children's National Medical Center (institutional project management, data collection); Morella Menicucci, MD, CCRP, Children's National Medical Center (data collection); Fidel Ramos, BA, Children's National Medical Center (institutional project management, data collection); Aimee Labell, MS, RN, Phoenix Children's Hospital (institutional project management, data collection); Courtney Bliss, BS, DTR, Phoenix Children's Hospital (data collection); Jeffrey Terry, MBA, Children's Hospital Los Angeles (data collection); Margaret Villa, RN, Children's Hospital Los Angeles and Mattel Children's Hospital UCLA (institutional project management, data collection); Jeni Kwok, JD, Children's Hospital Los Angeles and Mattel Children's Hospital (institutional project management, data collection); Amy Yamakawa, BS, Children's Hospital Los Angeles and Mattel Children's Hospital UCLA (data collection); Ann Pawluszka, BSN, RN, Children's Hospital of Michigan (institutional project management); Symone Coleman, BS, MPH, Children's Hospital of Michigan (data collection); Melanie Lulic, BS, Children's Hospital of Michigan (data collection); Mary Ann DiLiberto, BS, RN, CCRC, Children's Hospital of Philadelphia (institutional project management, data collection); Carolann Twelves, BSN, RN, Children's Hospital of Philadelphia (data collection); Monica S. Weber, RN, BSN, CCRP, University of Michigan (institutional project management, data collection); Lauren Conlin, BSN, RN, CCRP, University of Michigan (data collection); Alan C. Abraham, BA, CCRC, Children's Hospital of Pittsburgh of University of Pittsburgh Medical Center (institutional project management, data collection); Jennifer Jones, RN, Children's Hospital of Pittsburgh of University of Pittsburgh Medical Center (data collection); Jeri Burr, MS, RN-BC, CCRC, University of Utah (project management, Data Coordinating Center); Nichol Nunn, BS, MBA, University of Utah (project management, Data Coordinating Center); Alecia Peterson, BS, CMC, University of Utah (project management, Data Coordinating Center); Carol Nicholson, MD (former Project Officer, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, for part of the study period); Christopher J. L. Newth, MD, FRCPC, Department of Anesthesiology and Critical Care Medicine, Children's Hospital Los Angeles (site co-principal investigator), Los Angeles, Calif; Thomas Shanley, MD, Department of Pediatrics, University of Michigan (site principal investigator); Rick E. Harrison, MD, Department of Pediatrics, University of California at Los Angeles (site co-principal investigator); Allan Doctor, MD, Departments of Pediatrics and Biochemistry, Washington University School of Medicine (steering committee chair); Tammara L. Jenkins, MSN, RN, Pediatric Trauma and Critical Illness Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institutes of Health.

References

1. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg*. 2002;123:110-8.
2. O'Brien SM, Jacobs JP, Pasquali SK, et al. The Society of Thoracic Surgeons Congenital Heart Surgery Database Mortality Risk Model: Part 1-statistical methodology. *Ann Thorac Surg*. 2015;100:1054-62.
3. Jacobs JP, O'Brien SM, Pasquali SK, et al. The Society of Thoracic Surgeons Congenital Heart Surgery Database Mortality Risk Model: Part 2-clinical application. *Ann Thorac Surg*. 2015;100:1063-8; discussion 1068-70.
4. Jacobs JP, O'Brien SM, Pasquali SK, et al. The importance of patient-specific preoperative factors: an analysis of the society of thoracic surgeons congenital heart surgery database. *Ann Thorac Surg*. 2014;98:1653-8; discussion 1658-9.
5. Jacobs ML, O'Brien SM, Jacobs JP, et al. An empirically based tool for analyzing morbidity associated with operations for congenital heart disease. *J Thorac Cardiovasc Surg*. 2013;145:1046-57.e1041.
6. Jacobs JP, Shahian DM, Prager RL, et al. Introduction to the STS National Database Series: outcomes analysis, quality improvement, and patient safety. *Ann Thorac Surg*. 2015;100:1992-2000.
7. Pollack MM, Ruttimann UE, Getson PR. Pediatric Risk of Mortality (PRISM) score. *Crit Care Med*. 1988;16:1110-6.
8. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated Pediatric Risk of Mortality score. *Crit Care Med*. 1996;24:743-52.
9. Knaus WA, Zimmerman JE, Wagner DP, Draper EA, Lawrence DE. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med*. 1981;9:591-7.
10. Richardson DK, Gray JE, McCormick MC, Workman K, Goldmann DA. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. *Pediatrics*. 1993;91:617-23.

11. Averill RFGN, Hughes JS, Bonazelli J, McCullough EC, Steinbeck BA, Mullin R, et al. *All Patient Refined Diagnostic Related Groups (APR-DRG), version 20.0, methodology overview*. Wallingford: 3M Health Information Systems; 2003.
12. Pollack MM, Holubkov R, Funai T, et al. The Pediatric Risk of Mortality Score: update 2015. *Pediatr Crit Care Med*. 2016;17:2-9.
13. Pollack MM, Holubkov R, Funai T, et al. Simultaneous prediction of new morbidity, mortality, and survival without new morbidity from pediatric intensive care: a new paradigm for outcomes assessment. *Crit Care Med*. 2015;43:1699-709.
14. Jeffries HE, Soto-Campos G, Katch A, Gall C, Rice TB, Wetzel R. Pediatric index of Cardiac Surgical Intensive Care Mortality Risk Score for pediatric cardiac critical care. *Pediatr Crit Care Med*. 2015;16:846-52.
15. Czaja AS, Scanlon MC, Kuhn EM, Jeffries HE. Performance of the Pediatric Index of Mortality 2 for pediatric cardiac surgery patients. *Pediatr Crit Care Med*. 2011;12:184-9.
16. Straney L, Clements A, Parslow RC, et al. Paediatric index of mortality 3: an updated model for predicting mortality in pediatric intensive care. *Pediatr Crit Care Med*. 2013;14:673-81.
17. Thiagarajan RR, Nathan M. Pediatric Index of Cardiac Surgical Intensive Care Mortality: a new severity of illness score for cardiac surgical patients in ICUs. *Pediatr Crit Care Med*. 2015;16:885-6.
18. 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;138:1139-53.
19. Pollack MM, Holubkov R, Funai T, Clark A, Berger JT, Meert K, et al. Pediatric intensive care outcomes: development of new morbidities during pediatric critical care. *Pediatr Crit Care Med*. 2014;15:821-7.
20. Pollack MM, Holubkov R, Glass P, et al. Functional Status Scale: new pediatric outcome measure. *Pediatrics*. 2009;124:e18-28.
21. Pollack MM, Holubkov R, Funai T, et al. Relationship between the functional status scale and the pediatric overall performance category and pediatric cerebral performance category scales. *JAMA Pediatr*. 2014;168:671-6.
22. Fiser DH. Assessing the outcome of pediatric intensive care. *J Pediatr*. 1992;121:69-74.
23. Meert KL, Keele L, Morrison W, et al. End-of-life practices among tertiary care PICUs in the United States: a multicenter study. *Pediatr Crit Care Med*. 2015;16:e231-8.
24. Berg RA, Nadkarni VM, Clark AE, et al. Incidence and outcomes of cardiopulmonary resuscitation in PICUs. *Crit Care Med*. 2016;44:798-808.
25. Jenkins KJ, Gauvreau K. Center-specific differences in mortality: preliminary analyses using the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method. *J Thorac Cardiovasc Surg*. 2002;124:97-104.
26. Fagerland MW, Hosmer DW, Bofin AM. Multinomial goodness-of-fit tests for logistic regression models. *Stat Med*. 2008;27:4238-53.
27. Mossman D. Three-way ROCs. *Med Decis Making*. 1999;19:78-89.
28. Pollack MM, Dean JM, Butler J, et al. The ideal time interval for critical care severity-of-illness assessment. *Pediatr Crit Care Med*. 2013;14:448-53.
29. Zeltser I, Jarvik GP, Bernbaum J, et al. Genetic factors are important determinants of neurodevelopmental outcome after repair of tetralogy of Fallot. *J Thorac Cardiovasc Surg*. 2008;135:91-7.
30. Gaynor JW, Ittenbach RF, Gerdes M, et al. Neurodevelopmental outcomes in preschool survivors of the Fontan procedure. *J Thorac Cardiovasc Surg*. 2014;147:1276-82; discussion 1282-3.e1275.
31. Gaynor JW, Wernovsky G, Jarvik GP, et al. Patient characteristics are important determinants of neurodevelopmental outcome at one year of age after neonatal and infant cardiac surgery. *J Thorac Cardiovasc Surg*. 2007;133:1344-53.

Key Words: severity of illness, congenital heart disease, pediatric heart surgery, pediatrics, outcome prediction, critical care, pediatric critical care, intensive care, pediatric intensive care, pediatric risk of mortality, quality, quality assessment, physiological status, morbidity

APPENDIX E1. DESCRIPTION AND DETAILS OF THE TRICHOTOMOUS OUTCOME PREDICTION IN CRITICAL CARE STUDY

This investigation was performed in the Collaborative Pediatric Critical Care Research Network of the Eunice Kennedy Shriver National Institute of Child Health and Human Development. Patients from newborn to age 18 years were randomly selected and stratified by hospital from December 4, 2011, to April 7, 2013. Patients from both general/medical and cardiac/cardiovascular pediatric intensive care units (PICUs) were included. Moribund patients (vital signs incompatible with life for the first 2 hours after PICU admission) were excluded. Only the first PICU admission during a hospitalization was included. The protocol was approved by all institutional review boards. Multiple publications concerning these data have occurred.[E1-E8]

Selected site and patient-level data are shown in Appendix Table 1.

Patients were selected using a simple randomization scheme to ensure that patients were randomly selected on days when the number of patients admitted to the intensive care unit was high. For each weekday/weekend of recruitment into the Trichotomous Outcome Prediction in Critical Care study, a random shuffling of the digits 0 to 9 was randomly generated by the Coordinating Center. The last digit of the medical record number for each eligible patient was then compared with the relevant sequence to determine the patients to be approached. For example, if the sequence for a particular day was 8 3 9 2 0 1 7 5 4 6, then any patients with the last digit of the medical record number being 8 would be enrolled first, followed by any patients with the last medical record number digit being 3, and so on, until the limit for that weekday/weekend was reached for the center.

Outcomes

Morbidity affecting a significant decrement in functional status was assessed with the Functional Status Scale (FSS) and was recorded for the preacute illness (baseline) and at hospital discharge.[E9] It was determined from the medical records and/or discussions with the health care providers. Newborns never achieving a stable baseline were assigned an FSS score of 6; this was operationalized by assigning a FSS of 6 to admissions to the study sites from 0 to 2 days of age and to transfers from another facility from 3 to 6 days of age. Baseline FSS scores were categorized as 6 to 7 (good), 8 to 9 (mildly abnormal), 10 to 15 (moderately abnormal), 16 to 21 (severely abnormal), and >21 (very severely abnormal).[E5] New morbidity was defined as an increase in the FSS score $\geq$3 points from baseline to hospital discharge; changes of this magnitude indicate very significant worsening of functioning.[E5,E6,E9] Morbidity occurs in essentially all ages and types of patients, in relatively equal proportions, and involves all FSS domains.[E6]

Measurement of Physiological Status

Physiological status was measured with the Pediatric Risk of Mortality (PRISM) score with a shortened time interval (2 hours before admission to 4 hours after admission for laboratory data and the first 4 hours of PICU care for other physiologic variables).[E1,E7,E10] Appendix Table 2 contains the PRISM score variables and the points for each physiologic derangement.

PRISM components may be separated into cardiovascular (heart rate, systolic blood pressure, and temperature), neurologic (pupillary reactivity and mental status), respiratory (arterial partial pressure of oxygen, pH, partial pressure of carbon dioxide, and total bicarbonate), chemical (glucose, potassium, blood urea nitrogen, and creatinine), and hematologic (white blood cell count, platelet count, prothrombin, and partial thromboplastin time) components for modeling. PRISM may also be separated into neurologic and nonneurologic categories for modeling.

Congenital Cardiac Conditions

The timing interval for assessing PRISM data was modified for cardiac patients younger than age 91 days because some institutions admit infants to the PICU before a cardiac intervention to optimize the clinical status but not for intensive care; in these cases, the postintervention period more accurately reflects intensive care.[E10] However, in other infants for whom the cardiac intervention is delayed after PICU admission, the intervention is a therapy required due to failed medical management of the acute condition; in these infants, the routine PRISM data collection time interval is an appropriate reflection of critical illness. Therefore, we identified infants for whom it would be more appropriate to use data from the 4 hours after the cardiac intervention (postintervention time interval) and those for whom using the admission time interval was more appropriate. We operationalized this decision on the conditions likely to present within the first 90 days, the time period when the vast majority of these conditions present. For newborns younger than age 24 hours at PICU admission, we used the admission time interval if the intervention was within the first 12 hours of PICU care. For all other patients admitted within the first 10 days of age, we used postintervention PRISM time interval if the intervention occurred in the first 10 days of PICU care. For children aged more than 10 days and younger than age 31 days at PICU admission, we used the postintervention PRISM time interval if the intervention occurred within the first 48 hours of PICU care. For infants 31 to 90 days of age at PICU admission, we used the postintervention time interval for all cardiac surgeries if the surgery was within 48 hours of admission and the admission time interval for all cardiac interventional catheterizations. Infants aged >90 days at admission had the routine PRISM time interval used.

E-References

E1. Pollack MM, Holubkov R, Funai T, Dean JM, Berger JT, Wessel DL, et al. The Pediatric Risk of Mortality Score: Update 2015. *Pediatr Crit Care Med*. 2016; 17:2-9.

E2. Berg RA, Nadkarni VM, Clark AE, Moler F, Meert K, Harrison RE, et al. Incidence and outcomes of cardiopulmonary resuscitation in PICUs. *Crit Care Med*. 2016;44:798-808.

E3. Pollack MM, Holubkov R, Funai T, Berger JT, Clark AE, Meert K, et al. simultaneous prediction of new morbidity, mortality, and survival without new morbidity from pediatric intensive care: a new paradigm for outcomes assessment. *Crit Care Med*. 2015;43:1699-709.

E4. Meert KL, Keele L, Morrison W, Berg RA, Dalton H, Newth CJ, et al. End-of-life practices among tertiary care PICUs in the United States: a multicenter study. *Pediatr Crit Care Med*. 2015;16:e231-8.

E5. Pollack MM, Holubkov R, Funai T, Moler F, Meert K, Harrison RE, et al. Relationship between the functional status scale and the pediatric overall performance category and pediatric cerebral performance category scales. *JAMA Pediatr*. 2014;168:671-6.

E6. Pollack MM, Holubkov R, Funai T, Clark A, Berger JT, Meert K, et al. Pediatric intensive care outcomes: development of new morbidities during pediatric critical care. *Pediatr Crit Care Med*. 2014;15:821-7.

E7. Pollack MM, Dean JM, Butler J, et al. The ideal time interval for critical care severity-of-illness assessment. *Pediatr Crit Care Med*. 2013;14: 448-53.

E8. Keele L, Meert KL, Berg RA, Holubkov R, Doctor A, Meert KL, et al. Limiting and withdrawing life support in the PICU: for whom are these options discussed? *Pediatr Crit Care Med*. 2016;17:110-20.

E9. Pollack MM, Holubkov R, Glass P, Dean JM, Meert KL, Zimmerman J, et al. Functional Status Scale: new pediatric outcome measure. *Pediatrics*. 2009;124: e18-28.

E10. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated Pediatric Risk of Mortality score. *Crit Care Med*. 1996;24:743-52.

APPENDIX TABLE 1. Selected patient and site characteristics

- Sample size
- Age
- Sex
- Insurance: commercial/government/other
- Race: white/black/other
- Primary system of dysfunction
  - Respiratory
  - Cardiovascular disease – acquired
  - Cardiovascular disease – congenital
  - Neurologic
  - Endocrine
  - Gastrointestinal
  - Hematological
  - Musculoskeletal
  - Renal
  - Miscellaneous
- Admitted for postintervention care: yes/no
- Pediatric intensive care unit admission status
  - Elective (scheduled)
  - Emergency (unscheduled)
- Pediatric intensive care unit admission status
  - Elective
  - Emergency
- Intervention category
  - None
  - Cardiac surgery
  - Interventional cardiac catheterization
  - Neurosurgery
  - Orthopedic
  - General surgery
  - Otolaryngology
  - Miscellaneous
- Admission source
  - Operating/intervention room or postanesthesia care unit
  - Inpatient unit from same hospital
  - Direct admission from referring hospital
  - Emergency department same hospital
- Cardiac arrest before pediatric intensive care unit admission: yes/no
- Functional Status Scale
- Pediatric Risk of Mortality score
  - Cardiovascular variables
  - Metabolic variables
  - Chemistry variables
  - Hematologic variables
  - Neurologic variables
- Length of stay
- Outcome at hospital discharge
  - New morbidity
  - Death
  - Intact survival

APPENDIX TABLE 2. Pediatric Risk of Mortality score variables and scores for physiologic derangements[E1,E10]

Variable	Score		
Cardiovascular and neurologic vital signs			
Systolic blood pressure (mm Hg)	Score = 3	Score = 7	
Neonate	40-55	<40	
Infant	45-65	<45	
Child	55-75	<55	
Adolescent	65-85	<65	
Temperature	Score = 3		
All ages	<33°C (91.4°F) or >40°C (104.0°F)		
Mental status	Score = 5		
All ages	Stupor/coma or GCS <8		
Heart rate (beats per minute)	Score = 3	Score = 4	
Neonate	215-225	>225	
Infant	215-225	>225	
Child	185-205	>205	
Adolescent	145-155	>155	
Pupillary reflexes	Score = 7	Score = 11	
All ages	1 fixed	Both fixed	
Acid-base, blood gases (all ages)			
Acidosis (pH or total carbon dioxide)	Score = 2	Score = 6	
pH	7.0-7.28	<7.0	
Carbon dioxide	5.0-16.9	<5	
Partial pressure of carbon dioxide (mm Hg)	Score = 1	Score = 3	
	50-75	>75	
Alkalosis: Total carbon dioxide (mmol/L)	Score = 4		
	>34		
Partial pressure of oxygen (mm Hg)	Score = 3	Score = 6	
	42-49	<42	
Chemistry tests			
Glucose	Score = 2		
All ages	>200 mg/dL or >11 mmol/L		
Potassium (mmol/L)	Score = 3		
All ages	>6.9		
Blood urea nitrogen	Score = 3		
Neonate	>11.9 mg/dL or >4.3 mmol/L		
All other ages	>14.9 mg/dL or >5.4 mmol/L		
Creatinine	Score = 2		
Neonate	>0.85 mg/dL or >75 μmol/L		
Infant	>0.90 mg/dL or >80 μmol/L		
Child	>0.90 mg/dL or >80 μmol/L		
Adolescent	>0.1.3 mg/dL or >115 μmol/L		
Hematology tests			
White blood cell count (cells/mm^3)	Score = 4		
All ages	<3000		
Platelet count ($\times$ 10^3 cells/mm^3)	Score = 2	Score = 4	Score = 5
All ages	100-200	50-99	<50
Prothrombin time	Score = 3		
Neonate	>22.0		
All other ages	>22.0		
Partial thromboplastin time	Score = 3		
Neonate	>85.0		
All other ages	>57.0		

GCS, Glasgow Coma Scale score.

TABLE E1. PRISM III score sampling intervals for cardiac patients receiving an intervention. The admission time interval refers to the period 2 hours before admission to 4 hours after admission for laboratory data and the first 4 hours of ICU care for other physiologic variables. The postintervention time interval refers to the first 4 hours of ICU care after a cardiac intervention (surgery or interventional catheterization, but not diagnostic catheterization[12,13])

Admission age	ICU length of stay before cardiac intervention	PRISM III data collection time period
<24 h	≤12 h	Admission
	12 h-10 d	Postintervention
24 h-10 d	0-10 d	Postintervention
	>10 d	Admission
11-30 d	≤48 h	Postintervention
	>48 h	Admission
31-90 d	48 h	Postintervention if cardiac surgery Admission if cardiac catheterization
	≥48 h	Admission
>90 d	All	Admission

PRISM, Pediatric Risk of Mortality; *ICU*, intensive care unit.

TABLE E2. Institutional characteristics

	Site								Overall	
	A	B	C	D	E	F	G	H		*P*
Characteristic	(n = 220)	(n = 350)	(n = 174)	(n = 161)	(n = 74)	(n = 148)	(n = 76)	(n = 347)	(N = 1550)	value
Total surgeries	223	360	178	165	76	154	81	359	1596	
Age (mo)	0.6 (0.2, 3.8)	0.5 (0.0, 4.2)	0.4 (0.1, 3.7)	0.4 (0.0, 2.1)	1.0 (0.3, 5.8)	0.6 (0.3, 4.0)	0.6 (0.1, 6.6)	0.7 (0.2, 4.4)	0.6 (0.1, 4.1)	.006*
Age										.065*
Neonates (0-30 d)	48 (21.8)	99 (28.3)	43 (24.7)	49 (30.4)	15 (20.3)	27 (18.2)	14 (18.4)	77 (22.2)	372 (24.0)	
Infants (31-90 d)	18 (8.2)	19 (5.4)	20 (11.5)	17 (10.6)	1 (1.4)	5 (3.4)	10 (13.2)	17 (4.9)	107 (6.9)	
Infants (91-365 d)	69 (31.4)	80 (22.9)	40 (23.0)	37 (23.0)	21 (28.4)	50 (33.8)	19 (25.0)	95 (27.4)	411 (26.5)	
Child (1-12 y)	63 (28.6)	118 (33.7)	58 (33.3)	49 (30.4)	25 (33.8)	54 (36.5)	23 (30.3)	132 (38.0)	522 (33.7)	
Adolescent (>12 y)	22 (10.0)	34 (9.7)	13 (7.5)	9 (5.6%)	12 (16.2)	12 (8.1)	10 (13.2)	26 (7.5)	138 (8.9)	
STAT										.819*
1	60 (27.3)	103 (29.4)	46 (26.4)	46 (28.6)	20 (27.0)	34 (23.0)	18 (23.7)	96 (27.7)	423 (27.3)	
2	70 (31.8)	89 (25.4)	62 (35.6)	57 (35.4)	31 (41.9)	63 (42.6)	24 (31.6)	117 (33.7)	513 (33.1)	
3	35 (15.9)	58 (16.6)	18 (10.3)	17 (10.6)	8 (10.8)	23 (15.5)	7 (9.2)	39 (11.2)	205 (13.2)	
4	41 (18.6)	76 (21.7)	29 (16.7)	30 (18.6)	13 (17.6)	21 (14.2)	20 (26.3)	78 (22.5)	308 (19.9)	
5	10 (4.5)	23 (6.6)	14 (8.0)	10 (6.2)	2 (2.7)	5 (3.4)	5 (6.6)	16 (4.6)	85 (5.5)	
Unclassified	4 (1.8)	1 (0.3)	5 (2.9)	1 (0.6)	0 (0.0)	2 (1.4)	2 (2.6)	1 (0.3)	16 (1.0)	
Outcome‡										.081*
Intact survival	200 (90.9)	335 (95.7)	160 (92.0)	143 (88.8)	70 (94.6)	134 (90.5)	66 (86.8)	318 (91.6)	1426 (92.0)	
New morbidity	11 (5.0)	6 (1.7)	6 (3.4)	12 (7.5)	4 (5.4)	14 (9.5)	6 (7.9)	15 (4.3)	74 (4.8)	
Death	9 (4.1)	9 (2.6)	8 (4.6)	6 (3.7)	0 (0.0)	0 (0.0)	4 (5.3)	14 (4.0)	50 (3.2)	
Intensive care unit length of stay (d)	4.1 (2.4, 9.0)	4.3 (2.0, 10.9)	3.8 (1.9, 9.9)	4.2 (2.0, 11.3)	2.7 (1.1, 6.9)	6.1 (4.1, 15.1)	6.5 (3.7, 9.2)	3.2 (2.0, 8.1)	4.1 (2.1, 10.1)	<.001*
Hospital length of stay (d)	8.4 (5.2, 20.5)	10.3 (6.3, 20.1)	8.6 (4.4, 20.8)	11.7 (5.2, 24.8)	5.3 (3.3, 13.3)	6.3 (4.3, 17.3)	7.9 (4.3, 26.1)	6.4 (4.0, 15.3)	8.3 (4.4, 19.1)	<.001*
Cyanotic/acyanotic										.030†
Acyanotic	134 (60.9)	195 (55.7)	84 (48.3)	95 (59.0)	44 (59.5)	96 (64.9)	35 (46.1)	188 (54.2)	871 (56.2)	
Cyanotic	86 (39.1)	155 (44.3)	90 (51.7)	66 (41.0)	30 (40.5)	52 (35.1)	41 (53.9)	159 (45.8)	679 (43.8)	
Ventricles										.086†
Double ventricle	174 (79.1)	274 (78.3)	121 (69.5)	124 (77.0)	59 (79.7)	124 (83.8)	63 (82.9)	260 (74.9)	1199 (77.4)	
Single ventricle	46 (20.9)	76 (21.7)	53 (30.5)	37 (23.0)	15 (20.3)	24 (16.2)	13 (17.1)	87 (25.1)	351 (22.6)	
Pediatric risk of mortality score	4.0 (2.0, 8.0)	7.0 (3.0, 10.0)	4.0 (1.0, 6.0)	5.0 (2.0, 11.0)	4.5 (2.0, 7.0)	4.0 (2.0, 7.0)	7.0 (4.5, 13.5)	5.0 (3.0, 8.0)	5.0 (2.0, 9.0)	<.001*

Values are presented as n, median (quartile 1, quartile 3), or n (%). *STAT*, Society for Thoracic Surgery Congenital Heart Surgery Database Mortality Risk. *Kruskal-Wallis test across sites. †Fisher exact test (Monte-Carlo approximation) across sites. ‡Outcome determined at hospital discharge.

TABLE E3. Assessment of the performance of the Pediatric Risk of Mortality (PRISM) score prediction models in infants aged ≤90 days and >90 days. The PRISM observation period for infants aged ≤90 days is predetermined based on the age at admission and the time to cardiac intervention[12,13]

		Survival-death model			Morbidity-intact survival-death model					
		Mortality			Morbidity			Mortality		
Variable	n	O	E	SMR (95% confidence interval)	O	E	SMR (95% confidence interval)	O	E	SMR (95% confidence interval)
Age ≤90 d	479	34	37.2	0.9 (0.6-1.3)	46	35.5	1.3 (0.9-1.7)	34	38.0	0.9 (0.6-1.3)
Age >90 d	1071	16	12.9	1.2 (0.7-2.0)	28	35.5	0.8 (0.5-1.1)	16	11.8	1.4 (0.8-2.2)

O, Observed; *E*, expected; *SMR*, standardized morbidity/mortality ratio.