Comparison of Severity Scoring Systems in a Pediatric Intensive Care Unit in India: A Single-Center Prospective, Observational Cohort Study

Vinayak K. Patki[1]   Sandeep Raina[2]   Jennifer V. Antin[2]

[1] Department of Pediatrics, Institute of Medical Science & Research, Vidyagiri, Satara, Maharashtra, India

[2] Department of Pediatrics, Wanless Hospital, Miraj, Sangli, Maharashtra, India

Address for correspondence Vinayak K. Patki, MB, DNB, FCCP, Department of Pediatrics, Institute of Medical Science & Research, Vidyagiri, Satara 415102, Maharashtra, India (e-mail: patkivinayak@gmail.com).

Abstract

Objective   The aim of this study is to validate Pediatric Risk of Mortality (PRISM) and Pediatric Index of Mortality (PIM) scoring systems in a pediatric intensive care unit (PICU) in India.

Design   This is a single-center prospective cohort study.

Setting   This study was conducted at an eight-bed PICU.

Methods   PRISM- and PIM-based predicted mortality rates were calculated and compared in 120 pediatric patients.

Results   Estimated mortality using PRISM (5.68%) and PIM (8.84%) was lower than observed (21.7%) mortality. PIM had slightly better power of calibration than PRISM. The discriminatory performance of both models was comparable.

Conclusion   Both models can be validated with suitable changes according to PICU settings of India.

Keywords
► PRISM
► PIM
► PICU
► critically ill children

Introduction

The recent history of pediatric intensive care has been marked by an increasingly sophisticated and concentrated approach. Pediatric intensive care units (PICUs) are now complex and costly. With their evolution, new medical and social pressures related to intensive care have simultaneously emerged.[1] Mortality reduction is undoubtedly the most fundamental aim of a PICU. This is achieved by intensively monitoring and treating critically ill patients considered at high risk for mortality. A physician's accuracy in estimating mortality risk for patients admitted to PICUs is largely subjective. A probability model predicting mortality risk is a rational and objective way to quantify severity.[2]

Severity of illness is reflected by the magnitude of comorbidities and physiologic disturbances in critically ill children. These disturbances are estimated by measuring how far apart the physiologic variables are from the normal range and by objective weighting of these variables to directly reflect their contribution to the mortality risk. Since the early 1980s, various scoring systems have been used in PICUs to evaluate severity of illness. These scoring systems assist in the prediction of patient mortality and allow for comparisons of standards.[3] The scoring systems are also important while conducting clinical trials to remove the bias by selecting patients with similar severity of illness. The principal scores that have been developed for the pediatric population are the Pediatric Risk of Mortality (PRISM) and Pediatric Index of Mortality (PIM), with their most recent versions being PRISM III and PIM-2.[4–8] These scores were developed by identifying variables relevant to mortality risk and scoring them after a multivariate statistical analysis by logistic regression.[9]

The PRISM score was published in 1988 by Pollack et al and exhibited an excellent discriminatory and predictive performance with 14 variables (namely, systolic blood pressure,

diastolic blood pressure, heart rate, respiratory rate, partial pressure of arterial oxygen/fraction of inspired oxygen ratio [Pao_2/Fio_2 ratio], partial pressure of arterial carbon dioxide [$Paco_2$], Glasgow coma scale, pupillary reaction, prothrombin time ratio [test/control], total bilirubin, serum potassium, serum calcium, blood glucose, and serum bicarbonate) collected at 24 hours of PICU admission. It is still the most widely known predictor of mortality score and is used in PICUs as well as in clinical trials as a standard prognostic score for the evaluation of disease severity in pediatric patients. A revised version of the PRISM score, PRISM III, has been available since 1996, which, according to its authors, offers better predictive capability. However, a considerable fee is charged for using it routinely, which has limited its use, even in developed countries.[5–8]

The PIM was developed in 1996 and is a simple model that consists of eight variables (namely, elective admission [yes/no], premorbid clinical condition, pupillary response, base excess in arterial blood, Pao_2/Fio_2 ratio at that time, systolic blood pressure, and mechanical ventilation at any time during first hour in PICU [yes/no]) measured at the time of admission to the PICU.[7,8] One advantage of the PIM over the PRISM is the fact that the PIM is based on just eight variables, all of which are collected at the time of admission, which facilitates data collection and avoids any impact on the results from 24 hours of intensive management strategies.[8] PICU scoring systems have to be validated for ICU settings in India, as various other factors such as different patient profiles, greater load of severity of illnesses, and difference in the quality of care affect the general outcome of patients with respect to survival or mortality. The Pediatric Logistic Organ Dysfunction (PELOD) score has recently been validated with good discrimination.[10] PELOD, being an outcome score, evaluates the clinical course of a patient explaining severity of organ dysfunction, whereas PRISM and PIM are prognostic scores that predict the risk of death.

The performance of the PRISM and PIM models has been compared several times by the authors who developed the scores themselves,[11–14] but have rarely been compared independently and are mostly studied for developed nation settings. Data from developing nations have conflicting results, such as underprediction of mortality, poor sensitivity of either one or two of the scoring systems, and different calibrating and discriminative ability.[15–19] In this study, we compared the discriminative ability and calibration of PRISM and PIM in a PICU in India.

Materials and Methods

This study was conducted in the PICU at Wanless Hospital, Miraj Medical Center, Miraj, Maharashtra, India. It is an 8-bed PICU within a 500-bed tertiary care center. It admits pediatric patients younger than 18 years. There are at least three resident doctors pursuing postgraduate studies under the supervision of two senior consultants on duty each day. This was a prospective observational cohort study from December 1, 2011, to November 30, 2012, including patients between 1 month and 18 years of age. In total, 132

consecutive cases admitted to the PICU were studied. Patients expiring or discharged before 24 hours of ICU admission were excluded from the study. Data for calculating scores and predictive outcomes were recorded prospectively and with the techniques set out for each score (PRISM, first 24 hours after admission; PIM, within 1 hour of admission to the PICU). The PRISM and PIM scores and respective mortality rates were calculated using the models available online.

Demographic data, including age at admission and sex, were collected to characterize the sample. The outcome for all cases was documented as discharge or death. Length of hospital stay in the unit was also recorded. Simple descriptive analysis was utilized for the groups and subgroups under study (mean, median, and standard deviation). The Flora "z" test was utilized to compare the general similarity between observed mortality and that estimated by the standardized mortality rate (SMR). For aptness of the two models, the Hosmer–Lemeshow goodness-of-fit test was employed to test the agreement between observed and expected mortality (calibration).[20] The capacity for discrimination between survivors and moribund patients was evaluated using the typical area under a receiver operating characteristic curve (ROC curve) and quantitative correlation between the results of the scores were analyzed using the Spearman test.

Data were analyzed using statistical software IBM SPSS version 19 (Armonk, New York, United States). The study was approved by the Committee for Ethics in Research at the Wanless hospital. The study incurred no additional risk to patients, and informed consent was obtained.

Results

A total of 132 patients were admitted to the PICU during the duration of study. Twelve patients were excluded due to discharge or expiry before 24 hours of admission. No patient was excluded due to lack of data. The male to female ratio of the patients studied was 1.8. There were 36 (30%) patients between 1 month and 12 months, 35 (29.1%) patients between 13 months and 59 months, and 49 (40.9%) patients older than 60 months. Mean duration of stay of the patients in PICU in our study was 5.98 days with a median of 5 days (range, 1.1–21 days). Twenty-six (21.7%) patients expired. There was a total of 10 (8.3%) cardiovascular cases with 1 (10%) death, 20 (16.7%) respiratory cases with 3 (15%) deaths, 32 (26.7%) neurological cases with 9 (28.1%) deaths, 28 (23.3%) sepsis cases with 10 (35.7%) deaths, 11 (9.16%) renal cases with 2 (18.18%) deaths, and 8 (6.6%) gastrointestinal cases with 1 (12.5%) death. Among 11 (9.16%) miscellaneous cases, there were 3 hematological cases, 3 endocrinological cases, 4 poisoning cases, and 1 dermatological case (►Table 1).

Estimated mortality using PRISM and PIM was 6.81 (5.68%) and 10.61 (8.84%), respectively. These values correspond with an SMR (95% confidence interval) of 1.428 (1.219–1.672) for PRISM and 1.254 (1.136–1.383) for PIM (►Table 2). When tested by Flora z test, these values were within the limits for not rejecting the null hypothesis (less than 1.96 and greater than −1.96). Both PRISM and PIM showed good calibration.

Table 1 Distribution according to primary system involved

System	Cases (N = 120), n (%)	Expired (N = 26), n (%)
Cardiovascular	10 (8.3%)	1 (10%)
Respiratory	20 (16.7%)	3 (15%)
Neurological	32 (26.7%)	9 (28.1%)
Sepsis	28 (23.3%)	10 (35.7%)
Renal	11 (9.16%)	2 (18.18%)
Gastrointestinal	8 (6.6%)	1 (12.5%)
Miscellaneous	11 (9.16%)	0

PIM ($\chi^2 = 5.315$, $p = 0.723$, d = 8) had slightly better power of calibration than PRISM ($\chi^2 = 3.895$, $p = 0.565$, d = 8). The discriminatory performance of the models, measured by area under the ROC curve, resulted in an area of 0.923 (0.861–0.964) for the PRISM and 0.946 (0.890–0.979) for the PIM (►**Table 3**). Findings were shown to have a good discriminatory performance between survivals and nonsurvivals. We observed that PIM has slightly better discriminatory power than PRISM (►**Fig. 1**).

The estimated probabilities of death reveal a positive and significant correlation between the PRISM and the PIM, with Spearman correlation coefficient being $r = 0.65$ ($p < 0.001$) (►**Tables 4** and **5**). Both PRISM and PIM underpredicted mortality.

Discussion

We compared the performance of PRISM and PIM in our PICU. There are few published studies comparing the two models in India and other developing nations.[10,11] Both the models underpredicted mortality like previous studies from developing countries.[21–25] The underprediction in our study could relate to the patient population, different resource allocation, different practices in the PICU, and lesser physical and human

Table 2 Comparison of PRISM and PIM for incidence of mortality

	PRISM	PIM
Mean of mortality risk, mean ± SD	5.705 ± 12.80	8.872 ± 15.98
Median of mortality risk, % (IQR)	1.55 (0.7–3.45)	2.00 (1.4–9.25)
Estimated mortality, n (%)	6.81 (5.68%)	10.61 (8.84%)
SMR (95% CI)	1.428 (1.219–1.672)	1.254 (1.136–1.383)
Flora z test	1.20	1.602

Abbreviations: CI, confidence interval; IQR, interquartile range; SD, standard deviation; SMR, standard mortality ratio.
aWhen tested with Flora z test, PRISM and PIM are within the limits for not rejecting the null hypothesis (1.96 and −1.96).

Table 3 Performance of PRISM and PIM scores

	PRISM	PIM
Hosmer–Lameshow goodness-of-fit test	$\chi^2 = 3.895$, $p = 0.565$	$\chi^2 = 5.315$, $p = 0.723$
Area under an ROC (95% CI)	0.923 (0.861–0.964)	0.946 (0.890–0.979)

Abbreviations: CI, confidence interval; ROC, receiver operating characteristic curve.

resources managing more severely diseased patients. The demographic variables of our population were not similar to the original validation of this model. In PICUs in the developed world, the majority of low-risk admissions are surgical patients.[1,2,18,19] In our PICU, there were no patients admitted for elective surgery.

Calibration evaluates how well the model classifies subjects into low-, medium-, and high-risk categories. The value of p was required to be greater than 0.05 for good calibration of the model. Both PRISM and PIM showed good calibration. PIM had slightly better power of calibration than PRISM. The results are supported by the findings of Martha et al,[16] Taori et al,[23] and Qureshi et al.[25] In our study, PRISM is better calibrated than PIM in the high-risk mortality group (i.e., >30–100).

A discriminatory power of 0.90 or more is considered excellent, 0.80–0.89 as good, and 0.70–0.79 as fair discriminatory performance by the scoring model.[26–28] The closer the ROC curve area is to 1.0, the better the prediction model.[29] The discriminatory power evaluated using the ROC curve showed excellent discrimination for both PRISM (area under curve [AUC], 0.923) and PIM (AUC, 0.946). We observed that PIM has slightly better discriminatory power than PRISM.

Although still debatable, both discrimination and calibration are important in validation of any generic scoring system.[28] Both functions gain importance in the respective objective for which the scoring system is used. Discrimination is important for distinguishing the outcome, either
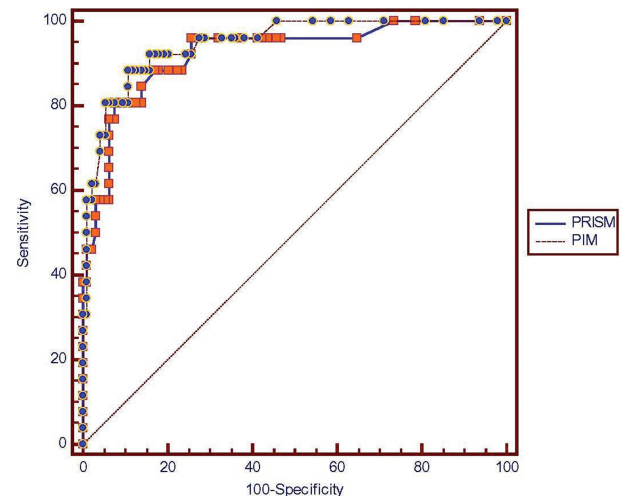

Fig. 1 Superposition of two receiver operating curves (ROC).

Table 4 Calibration of PRISM in five intervals of mortality risk

PRISM	Risk %	Number of patients	Observed survival	Expected survival	Observed death	Expected death
	0–1	51	50	50.68	1	0.32
	>1–5	45	38	43.97	7	1.03
	>5–15	14	06	12.66	8	1.34
	>15–30	5	0	3.92	5	1.08
	>30–100	5	0	1.96	5	3.04
	Total	120	94	113.19	26	6.81

survival or moribund, among the admitted patients. Calibration is more important for comparing expected and observed outcomes at various intervals of severity. Thus, discrimination and calibration are both important while validating prognostic scoring systems.[30] Our study demonstrated PIM as showing better calibration, though having only a slight edge over PRISM.

PIM had a much better discriminatory power compared with PRISM. Both PRISM and PIM showed good overall predictive performance. The results were in agreement with Taori et al,[23] Qureshi et al,[25] Martha et al,[16] and Slater and Shann,[12] who also showed good performance by PRISM and PIM in PICUs in developing countries. The estimated probabilities of death revealed a positive and significant correlation between the PRISM and the PIM, with Spearman correlation coefficient being $r = 0.65$ $(p < 0.001)$, which was similar to that observed by Martha et al,[16] but Qureshi et al[25] have reported higher correlation with $r = 0.74$ $(p < 0.001)$.

PIM has an edge over PRISM, as PIM has fewer variables thereby making assessment more convenient. As the resources are limited in developing countries such as India, this could make it economically more acceptable.

Comorbid conditions as well as diagnosis at admission to PICU invariably affect the outcome of patients. Wells et al[18] demonstrated a poor performance by the PRISM score, in terms of both calibration and discrimination. Poor performance could have been due to different demographic profile, disease distribution, or availability of infrastructure including trained personnel as well as equipment. Wells et al also attribute difficulty in achieving the same outcome for patients with a similar level of instability but having different pathological processes. PIM appears superior because the evaluation is done at the time of admission against 24 hours for PRISM. Early evaluation in PIM allows for commencing the intervention required in clinical trial early and effectively.

It is desirable that a scoring system should be devised that works in both developing and developed nations. This may involve modifying or adapting existing scoring systems in a way that may not affect their current functioning in the developed world but may appropriately modify their use in the developing world. The modification could take into account differences in the patient profile, differences in PICU practices, and differences in resource allocation.

Our study has a limitation of a small sample size compared with original validation studies. As such, we were not able to identify if the scoring system can predict mortality risk depending on primary pathology.

Conclusion

Both PRISM and PIM can be validated with suitable changes according to PICU settings of developing countries such as India. PIM has an edge over PRISM, as PIM has fewer variables, thus making assessment more convenient. Early evaluation with PIM allows for commencing the intervention required in clinical trial early and effectively. More information about the performance of the models in other regions of our country is required before these results can be generalized.

Table 5 Calibration of PIM in five intervals of mortality risk

PIM	Risk %	Number of patients	Observed survival	Expected survival	Observed death	Expected death
	0–1	6	6	5.96	0	0.04
	>1–5	75	73	73.61	2	1.33
	>5–15	21	13	19.01	8	1.99
	>15–30	9	1	6.98	8	2.02
	>30–100	9	1	3.77	8	5.23
	Total	120	94	109.39	26	10.61

References

1 Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score. Crit Care Med 1988;16(11):1110–1116

2 Seneff M, Knaus WA. Predicting patient outcome from intensive care: a guide to APACHE, MPM, SAPS, PRISM, and other prognostic scoring systems. J Intensive Care Med 1990;5(1):33–52

3 Teres D, Lemeshow S. Using severity measures to describe high performance intensive care units. Crit Care Clin 1993;9(3):543–554

4 Pollack MM, Yeh TS, Ruttiman UE, Holbrook PR, Fields AI. Evaluation of pediatric intensive care. Crit Care Med 1984;12(4):376–383

5 Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit Care Med 1996;24(5):743–752

6 Singhal D, Kumar N, Puliyel JM, Singh SK, Srinivas V. Prediction of mortality by application of PRISM score in intensive care unit. Indian Pediatr 2001;38(7):714–719

7 Shann F, Pearson G, Slater A, Wilkinson K. Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care. Intensive Care Med 1997;23(2):201–207

8 Lacroix J, Cotting J; Pediatric Acute Lung Injury and Sepsis Investigators (PALISI) Network. Severity of illness and organ dysfunction scoring in children. Pediatr Crit Care Med 2005;6(3, Suppl):S126–S134

9 Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982;115(1):92–106

10 Metta D, Soebardja D, Hudaya DS. The use of Pediatric Logistic Organ Dysfunction (PELOD) scoring system to determine the prognosis of patients in pediatric intensive care units. Paediatr Indones 2006;46:1–6

11 Pearson GA, Stickley J, Shann F. Calibration of the paediatric index of mortality in UK paediatric intensive care units. Arch Dis Child 2001;84(2):125–128

12 Slater A, Shann F; ANZICS Paediatric Study Group. The suitability of the Pediatric Index of Mortality (PIM), PIM2, the Pediatric Risk of Mortality (PRISM), and PRISM III for monitoring the quality of pediatric intensive care in Australia and New Zealand. Pediatr Crit Care Med 2004;5(5):447–454

13 van Keulen JG, Polderman KH, Gemke RJBJ. Reliability of PRISM and PIM scores in paediatric intensive care. Arch Dis Child 2005;90(2):211–214

14 Kapil D, Bagga A. The profile and outcome of patients admitted to a pediatric intensive care unit. Indian J Pediatr 1993;60(1):5–10

15 Thukral A, Lodha R, Irshad M, Arora NK. Performance of Pediatric Risk of Mortality (PRISM), Pediatric Index of Mortality (PIM), and PIM2 in a pediatric intensive care unit in a developing country. Pediatr Crit Care Med 2006;7(4):356–361

16 Martha VF, Garcia PC, Piva JP, Einloft PR, Bruno F, Rampon V. Comparison of two prognostic scores (PRISM and PIM) at a pediatric intensive care unit [in Portuguese]. J Pediatr (Rio J) 2005;81(3):259–264

17 Balakrishnan G, Aitchison T, Hallworth D, Morton NS. Prospective evaluation of the Paediatric Risk of Mortality (PRISM) score. Arch Dis Child 1992;67(2):196–200

18 Wells M, Riera-Fanego JF, Luyt DK, Dance M, Lipman J. Poor discriminatory performance of the Pediatric Risk of Mortality (PRISM) score in a South African intensive care unit. Crit Care Med 1996;24(9):1507–1513

19 Ozer EA, Kizilgunesler A, Sarioglu B, Halicioglu O, Sutcuoglu S, Yaprak I. The Comparison of PRISM and PIM scoring systems for mortality risk in infantile intensive care. J Trop Pediatr 2004;50(6):334–338

20 Hosmer DW, Lemeshow S. Applied Logistic Regression. New York, NY: John Wiley; 1989

21 Khilnani P, Sarma D, Zimmerman J. Epidemiology and peculiarities of pediatric multiple organ dysfunction syndrome in New Delhi, India. Intensive Care Med 2006;32(11):1856–1862

22 Bhal S, Tygai V, Kumar N, Sreenivas V, Puliyel JM. Signs of inflammation in children that can kill (SICK score): preliminary prospective validation of a new non-invasive measure of severity-of-illness. J Postgrad Med 2006;52(2):102–105

23 Taori RN, Lahiri KR, Tullu MS. Performance of PRISM (Pediatric Risk of Mortality) score and PIM (Pediatric Index of Mortality) score in a tertiary care pediatric ICU. Indian J Pediatr 2010;77(3):267–271

24 Choi KMS, Ng DKK, Wong SF, et al. Assessment of the Pediatric Index of Mortality (PIM) and the Pediatric Risk of Mortality (PRISM) III score for prediction of mortality in a paediatric intensive care unit in Hong Kong. Hong Kong Med J 2005;11(2):97–103

25 Qureshi AU, Ali AS, Ahmad TM. Comparison of three prognostic scores (PRISM, PELOD and PIM 2) at pediatric intensive care unit under Pakistani circumstances. J Ayub Med Coll Abbottabad 2007;19(2):49–53

26 Yeh TS, Pollack MM, Holbrook PR, Fields AI, Ruttiman U. Assessment of pediatric intensive care—application of the Therapeutic Intervention Scoring System. Crit Care Med 1982;10(8):497–500

27 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143(1):29–36

28 Murphy-Filkins R, Teres D, Lemeshow S, Hosmer DW. Effect of changing patient mix on the performance of an intensive care unit severity-of-illness model: how to distinguish a general from a specialty intensive care unit. Crit Care Med 1996;24(12):1968–1973

29 Te CT. Introductory Biostatistics. New Jersey, NJ: Wiley-Interscience; 2003

30 Mourouga P, Goldfrad C, Rowan KM. Does it fit? Assessment of scoring systems. Curr Opin Crit Care 2000;6(3):176–180