Comparison of the pediatric risk of mortality, pediatric index of mortality, and pediatric index of mortality 2 models in a pediatric intensive care unit in China
A validation study

Jun Qiu, PhD, Xiuulan Lu, MD, Kewei Wang, MD, Yimin Zhu, PhD, Chao Zuo, MD, Zhenghui Xiao, PhD*

Abstract
This study was designed with the aim of comparing the performances of the pediatric risk of mortality (PRISM), pediatric index of mortality (PIM), and revised version pediatric index of mortality 2 (PIM2) models in a pediatric intensive care unit (PICU) in China.

A total of 852 critically ill pediatric patients were recruited in the study between January 1 and December 31, 2014. The variables required to calculate PRISM, PIM, and PIM2 were collected. Model I performance was evaluated by assessing the calibration and discrimination. Discrimination between death and survival was assessed by calculating the area under the receiver-operating characteristic curve (AUC). Calibration across deciles of risk was evaluated using the Hosmer–Lemeshow goodness-of-fit $\chi^2$ test.

Of the 852 patients enrolled in this study, 745 patients survived until the end of the PICU stay (107 patients died, 12.56%). The AUCs (95% confidence intervals, CI) were 0.729 (0.670–0.788) for PRISM, 0.721 (0.667–0.776) for PIM, and 0.726 (0.671–0.781) for PIM2. The Hosmer–Lemeshow test revealed a chi-square of 7.26 ($P = 0.51, v = 10$) for PRISM, 26.28 ($P = 0.0009, v = 10$) for PIM, and 10.28 ($P = 0.21, v = 10$) for PIM2. The standardized mortality rate was 1.14 (95%CI: 0.93–1.36) for PRISM, 1.89 (95%CI: 1.55–2.27) for PIM, and 2.13 (95%CI: 1.75–2.55) for PIM2.

The PRISM, PIM, and PIM2 scores demonstrated an acceptable discriminatory performance. With the exception of PIM, the PRISM and PIM2 models had good calibrations.

Abbreviations: AUC = area under the receiver-operating characteristic curve, CI = confidence interval, ICU = intensive care unit, IQR = interquartile range, PICU = pediatric intensive care unit, PIM = pediatric index of mortality, PIM2 = pediatric index of mortality 2, PRISM = pediatric risk of mortality, SRR = standardized mortality ratio.

Keywords: pediatric index of mortality, pediatric intensive care unit, pediatric risk of mortality, prediction models

1. Introduction
The aim of pediatric intensive care units (PICUs) is to provide quality care for critically ill children. Mortality is the most frequently assessed outcome in modern intensive care units (ICUs) and is considered an important assessment criterion for medical services. However, similar conditions should be taken into account when comparing the quality and efficacy of medical services across different units and countries, such as severity of illness on admission and availability of medical resources. Therefore, 2 alternative methods of mortality prediction – the pediatric risk of mortality (PRISM and PRISM III)[1,2] and the pediatric index of mortality and pediatric index of mortality 2 (PIM and PIM2)[3,4] are used to measure the severity illness and the risk of mortality in PICUs. These models predict the risk of mortality by using the logistic regression modeling to obtain an equation that describes the relationship between predictor variables and the probability of death.

PRISM was developed using data collected from PICUs in the United States and was published in 1988 by Pollack et al.[1] It is a commonly used mortality prediction model, initially derived from the physiology stability index.[5] PRISM III, a 3rd-generation physiology-based prediction model for mortality, has been available since 1996 and offers an improved predictive capability.[2,6] PRISM has been widely used in both developed and developing countries, and the model has both satisfactory
discrimination and calibration. However, PRISM III has limitations, chiefly for its considerable price even in developed countries. And for this reason, it was not evaluated in this study.

PRISM was developed with data collected from both Australia and the United Kingdom between 1994 and 1996. It is a simple model consisting of 8 variables measured at the time of admission to an ICU. In 2003, the PRISM Study Group published a revised version of PRISM in Australia, New Zealand, and the United Kingdom. PRISM2 boasts its improved calibration, safety, and adjustment for various diagnoses when compared to the original version.

So far, the performance of the PRISM and PIM scoring systems has been compared only in certain specific disease categories or within heterogeneous groups of patients from PICUs. In China, the performance of these scores has been thoroughly investigated in Chinese neonatal ICUs. In 2015, our team has compared the performance of the PRISM and PIM2 models in Chinese PICUs. However, only 412 discontinuous critically ill pediatric patients transferred to Hunan Children’s Hospital were included, and the purpose of the study was to assess the performance of PRISM and PIM2 scores in predicting mortality in this patient population. Therefore, the present study was designed to compare the performances of PRISM, PIM, and PIM2 for 852 continuous critically ill pediatric patients in a PICU in China.

2. Materials and methods

2.1. Ethics statement

This study was approved by the Ethics Committee of Hunan Children’s Hospital with the trial registration number HCHLL-2014004. Informed written consent was obtained from the legal guardians of each child enrolled in this study prior to data collection. All data collected from the participants were fully anonymized.

2.2. Study subjects

All patients consecutively admitted to the PICU in Hunan Children’s Hospital between January 1 and December 31, 2014 were included in this study. Patients were excluded from this study if they died within 2 hours of admission or if they were discharged within 24 hours of admission. Hunan Children’s Hospital is the largest tertiary hospital in China, serving a population of 71 million and a land area of 211,800 km². The hospital has 1800 beds, and 40 of these are in the PICU. Hunan Province is located in central China and had a population of 222 million children in 2013.

Each patient received a chart recording age, sex, diagnosis, temperature, respiratory rate, heart rate, blood pressure, oxygen saturation, white blood cell count, platelet count, liver enzymes, prothrombin time, partial thromboplastin time, levels of bilirubin, glucose, blood urea, creatinine, sodium, potassium, and calcium, arterial blood gas, Glasgow Coma Score, pupillary reactions to bright light, high or low risk diagnosis, occurrence of mechanical ventilation at any time during the first hour in the ICU, elective admission to the ICU, recovery from surgery or a procedure as the main reason for the ICU admission, and admission following cardiac bypass. Patient outcomes (death or survival) and length of stay in the PICU were also recorded. All patients were classified on admission according to their diagnostic group as follows: central nervous system, respiratory, cardiovascular, hematological, and miscellaneous. The variables for PIM and PIM2 were collected between the 1st contact with a PICU doctor and up to the 1st hour after admission to the PICU; PRISM with the 14 physiologic values can be collected during the 1st 24 hours after admission to the PICU. Respiratory rate, heart rate, blood pressure (systolic/diastolic), PaO2/FiO2, PaCO2, prothrombin time, partial thromboplastin time, total bilirubin, calcium, potassium, glucose, HCO3−, and pupillary reactions are the exact predictor variables for PRISM. Elective admission, underlying condition, response of the pupils to bright light, mechanical ventilation, systolic blood pressure, base excess, and FiO2/PaO2 are the exact predictor variables for PIM. Elective admission, recovery postprocedure, cardiac bypass, high risk diagnosis, low risk diagnosis, no response of the pupils to bright light, mechanical ventilation, systolic blood pressure, base excess, and FiO2/PaO2 are the exact predictor variables for PIM2. The PRISM, PIM, and PIM2 scores were calculated using the formula published in the original articles. Data were collected by a team of experienced research nurses, and the treating team was blinded to these scores and predictions. The probability of mortality was only calculated at the end of the study period to avoid bias in patient management.

2.3. Estimations of sample size

This is a validation study comparing 3 prognostic scores (PRISM, PIM, and PIM2) in a PICU. We used the sample size calculating formula for a validation study, n = \( \left( \frac{Z_{1-\alpha/2} + Z_{1-\beta}}{\delta} \right)^2 \)
, where \( Z_{1-\alpha/2} \) and \( Z_{1-\beta} \) are the standard normal deviates with a 0.05 level of significance, \( \beta \) is 0.8, and \( \alpha \) is the area under the receiver-operating characteristic curve. The estimated sample size was 779.

2.4. Statistical analysis

All analyses were performed using the Statistical Package for the Social Sciences (SPSS, Windows version 18.0) and Stata (version 7.0) software packages. Demographics and lengths of hospital and PICU stay were presented using median and interquartile ranges (IQRs). Categorical variables were compared using the chi-square test, while medians of continuous variables were compared using either the Mann–Whitney or Kruskal–Wallis tests. Model performance was evaluated by assessing the calibration and discrimination. Discrimination between death and survival was assessed by calculating the area under the receiver-operating characteristic curve (AUC). An AUC > 0.70 was considered an acceptable discriminatory performance. Calibration across deciles of risk was evaluated using the Hosmer–Lemeshow goodness-of-fit \( \chi^2 \) test; this test was calculated as \( \sum (O - E)^2/E \), where O is the observed and E the expected number of events in each group of risk. For this test, a P-value > 0.05 indicates good calibration. The overall performance of the 3 scoring systems was assessed using a standardized mortality ratio (SMR), which is the ratio of the risk-adjusted observed mortality to the expected mortality derived from the development set where the score was developed. Calibration across the diagnosis was assessed by calculating the SMR with a 95% confidence intervals (CIs) assessed using the Fisher exact test. Pearson correlation analysis was applied to the estimated probabilities of death to show the correlations among the PRISM,
Table 1
Characteristics of the pediatric patients in PICU.

| Variable            | All patients (n = 852) | Survivors (n = 745) | Deaths (n = 107) | (Z/χ²) | P   |
|---------------------|------------------------|---------------------|------------------|--------|-----|
| Age, months         | 6.5 (2–21)             | 6 (2–20.5)          | 8 (2–27)         | −1.32  | >0.05 |
| Range               | 1–204                  | 1–204               | 1–139            |        |     |
| Age categories, n, %|                        |                     |                  |        |     |
| <12 months          | 528 (61.07)            | 466 (62.55)         | 62 (57.94)       | 4.66   | >0.05 |
| 12–60 months        | 229 (26.88)            | 192 (25.77)         | 37 (34.58)       |        |     |
| 60–120 months       | 64 (7.51)              | 58 (7.79)           | 6 (6.61)         |        |     |
| >120 months         | 31 (3.64)              | 29 (3.89)           | 2 (1.87)         |        |     |
| Sex                 |                        |                     |                  |        |     |
| Male n, %           | 518 (60.80)            | 458 (61.48)         | 60 (56.07)       | 1.15   | >0.05 |
| Female n, %         | 334 (39.20)            | 287 (38.52)         | 47 (43.93)       |        |     |
| Duration of stay, P30 (P25–P75) | 8 (4–15)          | 9 (4–16)            | 3 (1–8)         | −7.01  | <0.05 |
| Admission type n, % |                        |                     |                  |        |     |
| Elective            | 13 (1.53)              | 13 (1.74)           | 0 (0)            | 1.89   | >0.05 |
| Urgent              | 839 (98.47)            | 732 (98.26)         | 107 (100)        |        |     |
| Mechanical ventilation in 1st hour, n, % |                    |                     |                  |        |     |
| Yes                 | 179 (21.31)            | 131 (17.87)         | 48 (44.86)       | 40.56  | <0.05 |
| No                  | 661 (78.69)            | 602 (82.13)         | 59 (55.14)       |        |     |
| Diagnosis n, %      |                        |                     |                  |        |     |
| Respiratory         | 353 (41.43)            | 332 (40.58)         | 21 (59.50)       | 35.61  | 0.001|
| Neurological        | 103 (12.09)            | 83 (10.00)          | 20 (19.42)       |        |     |
| Cardiac             | 26 (3.05)              | 22 (2.78)           | 4 (15.38)        |        |     |
| Injury              | 52 (6.12)              | 47 (50.38)          | 5 (9.62)         |        |     |
| Miscellaneous       | 246 (28.87)            | 195 (23.72)         | 51 (20.73)       |        |     |
| Postoperative       | 21 (2.46)              | 19 (20.48)          | 2 (0.95)         |        |     |
| Hematological       | 45 (5.28)              | 41 (48.46)          | 4 (8.89)         |        |     |
| Others              | 6 (0.70)               | 6 (100.00)          | 0 (0.00)         |        |     |
| PRISM, P30 (P25–P75) | 6.20 (3.06–12.25)      | 6.14 (2.81–10.76)   | 16.66 (6.81–42.73)| −7.67  | 0.001|
| PIM, P30 (P25–P75)  | 2.54 (1.90–6.92)       | 2.38 (1.67–5.16)    | 8.41 (2.59–19.58)| −7.41  | 0.001|
| PIM2, P30 (P25–P75) | 2.20 (1.59–5.78)       | 2.05 (1.54–4.82)    | 6.70 (2.22–15.46)| −7.57  | 0.001|

PICU = pediatric intensive care unit, PIM = pediatric index of mortality, PIM2 = pediatric index of mortality 2, PRISM = pediatric risk of mortality.

3. Results

A total of 885 patients were enrolled in this study between January 1 and December 31, 2014. However, 13 patients were discharged within 2 hours of admission for unknown reasons, 8 patients died within 2 hours of admission after cardiopulmonary resuscitation, and 12 patients were excluded due to lack of data. The final dataset for PRISM, PIM, and PIM2 analysis consisted of 852 patients with 107 (12.56%) recorded deaths. The median age was 6.5 months (IQR: 2–21), and 745 patients survived at the end of the PICU stay. Age and sex distribution were not significantly different between the survival and death groups (P > 0.05). The majority of the patients had respiratory diseases, nervous system diseases, and miscellaneous conditions. The median duration of hospital stay was 8 days (IQR: 4–15), and the length of hospital stay for the survival group was longer than that for the death group (Z = −7.01, P < 0.05). The PRISM median score was 11 (IQR: 7–14) for all patients, while that for the deaths was higher than that for the survival group (P < 0.05). The median probability of mortality for PRISM, PIM, and PIM2 scores was 6.20% (3.06%–12.23%), 2.54% (1.90%–6.92%), and 2.20% (1.59%–5.78%), respectively, and was significantly higher in the death group than in the survival group for the 3 scores (P < 0.05). Patient characteristics are presented in Table 1.

Table 2
Performance of the models.

|                | PRISM      | PIM        | PIM2       |
|----------------|------------|------------|------------|
| Mean of mortality risk, % (SD) | 11.05±13.94 | 6.63±11.60 | 5.89±11.10 |
| Median of mortality risk, % (IQR) | 6.20 (3.06–12.25) | 2.54 (1.90–6.93) | 2.20 (1.59–5.78) |
| Estimated mortality, % (IQR) | 94.14 (11.00%) | 56.50 (6.33%) | 50.20 (6.89%) |
| SMR (95% CI) | 1.14 (0.93–1.36) | 1.89 (1.55–2.27) | 2.13 (1.75–2.55) |
| Hosmer–Lemeshow goodness of fit test | χ² = 7.26, P = 0.51 | χ² = 26.28, P = 0.0009 | χ² = 10.28, P = 0.21 |
| Area under an ROC (95%CI) | 0.729 (0.670–0.780) | 0.721 (0.667–0.776) | 0.726 (0.671–0.781) |

CI = confidence interval, IQR = interquartile, PIM = pediatric index of mortality, PIM2 = pediatric index of mortality 2, PRISM = pediatric risk of mortality, ROC = receiver-operating characteristic, SD = standard deviation, SMR = standardized mortality ratio.
Expected mortality was 94.14 patients (11.05%) by PRISM, 56.50 (6.63%) by PIM, and 50.20 (5.89%) by PIM2. The SMR for the entire sample was 1.14 (95% CI: 0.93–1.36) for PRISM, 1.89 (95% CI: 1.55–2.27) for PIM, and 2.13 (95% CI: 1.75–2.53) for PIM2 in this study, respectively. The AUC (95% CI) was 0.729 (0.670–0.788) for PRISM, 0.721 (0.667–0.776) for PIM, and 0.726 (0.671–0.781) for PIM2 (Fig. 1).

Table 3 shows the model calibration across the various levels of probability of death. The PRISM score under-predicted mortality at 0 to 0.061978 and 0.252372 to 0.999999 predicted probability levels. The PIM and PIM2 2 scores under-predicted mortality at 10 predicted probability all levels. The Hosmer–Lemeshow test revealed a chi-square of 7.26 (P = 0.51, v = 10) for PRISM, 26.28 (P = 0.0009, v = 10) for PIM, and 10.78 (P = 0.21, v = 10) for PIM2.

The estimated probabilities of death revealed a positive and significant correlation between PRISM and the 2 PIM models with a Pearson correlation coefficient of r = 0.490 (P < 0.001) and 0.477 (P < 0.001), respectively. The correlation coefficient between PIM and PIM2 was 0.938 (P < 0.001) (Table 4). The Bland–Altman plot was used to reveal the differences between the 2 scores. The Bland–Altman plot with linear regression analyses with 95% confidence limits is presented in Figs. 2–4. A significant difference among the PRISM, PIM, and PIM2 was not observed.

### Table 3

| Probability of death, % | Survivors | Nonsurvivors | No. of patients | SMR |
|-------------------------|-----------|--------------|----------------|-----|
|                         | Observed  | Expected     | Observed       | Expected |
| PRISM 0.0000 to 0.023844|           | 81           | 83.69          | 4     | 85 | 3.05 (0.79–6.78) |
| 0.023844 to 0.027923   | 74        | 76.91        | 5              | 2.09   | 79 | 2.39 (0.75–4.95) |
| 0.027923 to 0.030622   | 79        | 83.37        | 5              | 2.68   | 86 | 2.61 (1.04–4.31) |
| 0.030622 to 0.050738   | 76        | 81.36        | 9              | 3.64   | 85 | 1.12 (4.35–2.47) |
| 0.050738 to 0.061978   | 73        | 74.44        | 6              | 4.56   | 79 | 1.32 (0.47–2.58) |
| 0.061978 to 0.0808024  | 81        | 79.72        | 5              | 6.28   | 86 | 0.79 (0.25–1.65) |
| 0.0808024 to 0.107551  | 82        | 78.05        | 4              | 7.95   | 86 | 0.50 (0.13–1.12) |
| 0.107551 to 0.146540   | 80        | 79.94        | 6              | 10.06  | 86 | 0.60 (0.13–1.12) |
| 0.146540 to 0.252372   | 70        | 69.98        | 15             | 15.02  | 85 | 0.90 (0.56–1.57) |
| 0.252372 to 0.9999999  | 49        | 54.45        | 46             | 40.55  | 95 | 1.13 (0.83–1.49) |
| Total                  | 745       | 757.86       | 107            | 94.14  | 852| 1.14 (0.93–1.36) |

| PIM² | Observed  | Expected |
|------|-----------|----------|
| 0.0000 to 0.015349 | 81 | 84.02 |
| 0.015349 to 0.018061 | 78 | 83.56 |
| 0.018061 to 0.020206 | 82 | 83.38 |
| 0.020206 to 0.023237 | 81 | 83.19 |
| 0.023237 to 0.025541 | 80 | 82.90 |
| 0.025541 to 0.030832 | 73 | 82.65 |
| 0.030832 to 0.047549 | 78 | 81.88 |
| 0.047549 to 0.083676 | 75 | 79.38 |
| 0.083676 to 0.131215 | 67 | 76.43 |
| 0.131215 to 0.9999999 | 56 | 56.19 |
| Total | 745 | 759.5 |

| PIM² | Observed  |
|------|-----------|
| 0.0000 to 0.017220 | 83 | 85.25 |
| 0.017220 to 0.014702 | 78 | 83.84 |
| 0.014702 to 0.016848 | 81 | 83.65 |
| 0.016848 to 0.019001 | 81 | 83.48 |
| 0.019001 to 0.021903 | 78 | 83.28 |
| 0.021903 to 0.023759 | 76 | 82.91 |
| 0.023759 to 0.045514 | 74 | 82.05 |
| 0.045514 to 0.069251 | 73 | 80.13 |
| 0.069251 to 0.111033 | 70 | 77.70 |
| 0.111033 to 0.9999999 | 51 | 59.51 |
| Total | 745 | 801.80 |

| χ² = 7.36, P = 0.51 |
| χ² = 26.28, P = 0.0009 |
| χ² = 10.78, P = 0.21 |

PIM = pediatric index of mortality; PIM² = pediatric index of mortality 2; PRISM = pediatric risk of mortality; ROC = receiver-operating characteristic; SMR = standardized mortality ratio.
4. Discussion

In this study, we compared the performance of 3 major published mortality predicting tools in PICUs in China. The major findings were as follows: all 3 scoring systems demonstrated acceptable discrimination between death and survival with AUCs > 0.70; the fit between observed and expected outcomes was close for all 3 models according to the SMR in this setting; and predicted results were similar to those observed for PRISM and PIM2 in terms of evaluating calibration by the Hosmer–Lemeshow goodness-of-fit test (with the exception of PIM which displayed poor calibration). The performances of PRISM, PIM, and PIM2 have thus far been evaluated in only 1 study in China. The study was performed on 243 consecutive neonate admissions in a Chinese neonatal ICU and demonstrated good discrimination and calibration in all 3 models.\[^{[17]}\] It further showed that discrimination in the 3 models was similar to certain other publications\[^{[12,24]}\] but was slightly weaker than that of several studies conducted in both developed\[^{[13,25]}\] and developing countries.\[^{[14,26]}\] In a study conducted in India, Singhal et al\[^{[27]}\] demonstrated that PRISM had an acceptable predictive value (AUC = 0.72). However, in a study conducted in The Netherlands, Visser et al\[^{[8]}\] reported that the AUC values for PIM, PIM2, PIM2-ANZ06, PIM2-ANZ08, and PRISM were > 0.8 in 8 PICUs. In a study conducted in Iran, the AUCs for PRISM were reported to be 0.803.\[^{[27]}\] This discrepancy may be attributed to the variances in the characteristics of the study populations, small sample sizes (particularly for number of recorded deaths), and improvement in the quality of medical care in the past several decades since the development of these predictive mortality models.

In the present study, observed mortality was significantly higher than that predicted (SMR > 1) by the PRISM, PIM, and PIM2. Our findings, however, differ from those of the majority of studies conducted in developing or developed countries, which reported these models to be either under- or over-predicted the mortality rates in their samples.\[^{[14–16,23,28,29]}\] For example, studies conducted in both Hong Kong and India have reported SMRs ranging from 0.61 to 3.3 for PIM and from 0.79 to 1.20 for PRISM.\[^{[14–16]}\] In Pakistan, a study has demonstrated that the SMRs (95% CI) using the PRISM and PIM2 models were 1.20 (0.94–1.50) and 1.57 (1.24–1.59), respectively.\[^{[27]}\] The authors have attributed this to the variances in patient profiles, medical resources, and quality of intensive care in PICUs. The variation in SMR may be attributed to the threshold for initiating, timing of intensive care and quality of care, and the accuracy of data collection.

In the present study, under-prediction in certain predicted probability levels could be related to the PICU in question. First, the majority of patients enrolled in this study had respiratory disorders. However, patients admitted to the PICU with respiratory dysfunction would have been treated at either the emergency unit or the referring hospital and may therefore have blood gas analysis results within normal limits. And patients with mechanical ventilation have normal blood gases at admission to PICU, which lead to the lower value of FiO2/PaO2 and base

| Table 4 | The correlation relationship among PRISM, PIM, and PIM2. |
|---------|----------------------------------|
|         | PRISM  | PIM  | PIM2 |
| PRISM   | 1.00   |      |      |
| PIM     | 0.490\(^*\) | 1.00 |      |
| PIM2    | 0.477\(^*\) | 0.938\(^*\) | 1.00 |

\(^*\) P < 0.05.
Several other studies have also shown good discrimination but

tively. A study conducted by Martha et al[10] has demonstrated that
different intervals of severity, then calibration capacity is of
then the capacity to discriminate is the most important. However,
distinguish between survival and nonsurvival among patients,

Chinese hospital setting.

Further con

in other developing countries.[10,27]

scoring model.[31] The importance of either function is dependent
and calibration are important in the validation of any prognostic
models and compared them for both discrimination
and PIM2 models and compared them for both discrimination
and calibration. Although still debatable, both discrimination
and calibration are important in the validation of any prognostic
scoring model.[51] The importance of either function is dependent
on the objective of prognostic score use.[52] If the study aims to
distinguish between survival and nonsurvival among patients,
then the capacity to discriminate is the most important. However,
if the study aims to compare observed and expected mortalities at
different intervals of severity, then calibration capacity is of
greater importance. Thus, discrimination and calibration are
both essential for global evaluation of the scores.

In our study, the correlation between PRISM and the 2 PIM
models was \( r = 0.490 \) (\( P < 0.001 \)) and 0.477 (\( P < 0.001 \)), respectively.
A study conducted by Martha et al[10] has demonstrated that
the correlation coefficient between PRISM and PIM models was
\( r = 0.65 \) (\( P < 0.05 \)); the correlation coefficient was lower than that in
other developing countries.[10,27]

Our study has demonstrated that the performance of the 3
models is similar in terms of their capacity to discriminate
between surviving and moribund patients. Nevertheless, PIM
exhibited a poor calibration capacity, which was also observed by Bertolini et al.[25] The power of calibration was tested in all the
scoring systems using the Hosmer–Lemeshow goodness of fit test,
and a \( P > 0.05 \) was considered a good calibration for the model.
Several other studies have also shown good discrimination but
poor calibration for PRISM, PIM, and PIM2 models.[15,24,29,33]

This poor calibration may be attributed to factors such as poorly
performing healthcare systems and limited resources in develop-
ing countries. A further possibility is the low number of deaths
reported at each level, and a study by Pearson et al[34]
recommends that special care be taken when variance is low
within a small series (eg., <20 deaths per unit). Other possible
contributors include different case mixes,[15] disease patterns,[36]
and critically ill patients with moderate or severe malnourish-
ment.[14,28]

5. Limitations

A major limitation in this study was the small sample size when
compared to the original validation studies. Additionally, all
data used in this study were obtained from a single PICU. The
small sample size is likely to interfere with the accurate
application of the Hosmer–Lemeshow test for goodness of fit.
Further confirmation of the results obtained in this study is
warranted before the generalized use of these scores in a
Chinese hospital setting.

6. Conclusion

This study has established that PRISM, PIM, and PIM2 scores
discriminate between surviving and moribund patients. Both the
PRISM and PIM2 models displayed good calibrations, while this
was not the case for PIM. However, a positive correlation was
observed across the PRISM, PIM, and PIM2 scores. Therefore,
the application of PRISM had a good score in PICU in China.

Acknowledgments

The authors thank all the patients who volunteered to participate
in this study. The authors also thank to Fatch Kalembro for the
language input, and the support of the data collectors.

References

[1] Pollack MM, Ruttmann UE, Getson PR. Pediatric risk of mortality
(PRISM) score. Crit Care Med 1988;16:1110–6.
[2] Pollack MM, Patel KM, Ruttmann UE. PRISM III: an updated pediatric
risk of mortality score. Crit Care Med 1996;24:743–52.
[3] Shann F, Pearson G, Slater A, et al. Paediatric index of mortality (PIM): a
mortality prediction model for children in intensive care. Intensive Care
Med 1997;23:2017–7.
[4] Slater A, Shann F, Pearson G, PIM2: a revised version of the paediatric
index of mortality. Intensive Care Med 2003;29:278–85.
[5] Yeh TS, Pollack MM, Ruttmann UE, et al. Validation of a physiologic
stability index for use in critically ill infants and children. Pediatr Res
1984;18:445–51.
[6] Marcin JP, Pollack MM, Patel KM, et al. Combining physician’s
subjective and physiology-based objective mortality risk predictions. Crit
Care Med 2000;28:2984–90.
[7] Khajeh A, Nooni NM, Resi M, et al. Mortality risk prediction by
application of pediatric risk of mortality scoring system in pediatric
intensive care unit. Ir J Pediatr 2013;23:546–50.
[8] Visser IH, Hazelzet JA, Albers MJ, et al. Mortality prediction models for
pediatric intensive care: comparison of overall and subgroup specific
performance. Intensive Care Med 2013;39:942–50.
[9] Bellad R, Rao S, Patil VD, et al. Outcome of intensive care unit patients
using pediatric risk of mortality (PRISM) score. Indian Pediatr
2009;46:1091–2.
[10] Martha VF, Garcia PC, Piva JP, et al. Comparison of two prognostic
scores (PRISM and PIM) at a pediatric intensive care unit. J Pediatr (Rio
J) 2005;81:239–64.
[11] Slater A. Monitoring outcome in paediatric intensive care. Paediatr
Anaesth 2004;14:113–6.
[12] Sankar J, Singh A, Sankar MJ, et al. Pediatric index of mortality and
PIM2 scores have good calibration in a large cohort of children from a
developing country. Biomed Res Int 2014;2014:907871.
[13] Slater A, Shann F. The suitability of the pediatric index of mortality
(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;5:447–54.
[14] Thukral A, Lodha R, Irshad M, et al. 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;7:356–61.
[15] Taori RN, Lahiri KR, Tuliu 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;77:267–71.
[16] Gandhi J, Sangareddi S, Varadarajan P, et al. Pediatric index of mortality
2 score as an outcome predictor in pediatric Intensive Care Unit in India.
Indian J Crit Care Med 2013;17:288–91.
[17] Wang C, Xie G, Cheng R, et al. Performance of pediatric risk of
mortality, pediatric index of mortality and PIM2 in term Chinese
neonates. J Trop Pediatr 2010;56:235–41.
[18] Lu X, Qiu J, Zhu Y, et al. [Performance of pediatric risk of mortality and
pediatric index of mortality in evaluation of the patients’ condition].
Zhonghua Er Ke Za Zhi 2015;53:370–4.
[19] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver
operating characteristic (ROC) curve. Radiology 1982;143:29–36.
[20] Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots:
a fundamental evaluation tool in clinical medicine. Clin Chem
1993;39:561–77.
[21] Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. Pediatr Crit Care Med 2005;6:2–8.
[22] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley; 1989:14–15, 141.
[23] Rapoport J, Teres D, Lemeshow S, et al. A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. Crit Care Med 1994;22:1385–91.
[24] degli Atti ML, Cuttini M, Rava L, et al. Performance of the pediatric index of mortality 2 (PIM-2) in cardiac and mixed intensive care units in a tertiary children’s referral hospital in Italy. BMC Pediatr 2013;13:100.
[25] Brady AR, Harrison D, Black S, et al. Assessment and optimization of mortality prediction tools for admissions to pediatric intensive care in the United Kingdom. Pediatrics 2006;117:e733–42.
[26] Choi KM, Ng DK, 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;11:97–103.
[27] Singhal D, Kumar N, Pujely JM, et al. Prediction of mortality by application of PRISM score in intensive care unit. Indian Pediatr 2001;38:714–9.
[28] 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;19:49–53.
[29] Bertolini G, Ripamonti D, Cattaneo A, et al. Pediatric risk of mortality: an assessment of its performance in a sample of 26 Italian intensive care units. Crit Care Med 1998;26:1427–32.
[30] Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 2007;35:2052–6.
[31] Randolph AG, Guyatt GH, Calvin JE, et al. Understanding articles describing clinical prediction tools. Evidence based medicine in critical care group. Crit Care Med 1998;26:1603–12.
[32] Mourouga P, Goldfrad C, Rowan KMD. Does it fit? Is it good? Assessment of scoring systems. Curr Opin Crit Care 2000;6:176–80.
[33] Canonero I, Figueroa A, Cacciamano A, et al. [Validation of PRISM and PIM2 scores of mortality in a pediatric intensive care unit in Cordoba]. Arch Argent Pediatr 2010;108:23–5.
[34] Pearson GA, Stickley J, Shann F. Calibration of the paediatric index of mortality in UK paediatric intensive care units. Arch Dis Child 2001;84:125–8.
[35] Pappachan JV, Millar B, Bennett ED, et al. Comparison of outcome from intensive care admission after adjustment for case mix by the APACHE III prognostic system. Chest 1999;115:802–10.
[36] Markgraf R, Deutschendorff G, Pietka L, et al. Comparison of acute physiology and chronic health evaluations II and III and simplified acute physiology score II: a prospective cohort study evaluating these methods to predict outcome in a German interdisciplinary intensive care unit. Crit Care Med 2000;28:26–33.